OXFORD ECONOMIC PAPERS

(NEW SERIES)

VOLUME 7 FEBRUARY 1955 NUMBER 1

Editorial Board: T. WILSON (*General Editor*), A. BRIGGS, J. R. HICKS
C. M. KENNEDY, SIR DONALD MACDOUGALL
G. B. RICHARDSON (*Secretary*)

CONTENTS

Some Implications of the Growth of Investment by Institutions, 1938–53. By R. FROST 1

The Rehabilitation of Monetary Theory in Sweden. By E. E. FLEETWOOD 11

The Intransitivity of Certain Criteria Used in Welfare Economics. By W. M. GORMAN 25

The Valuation of Net Investment. By C. KENNEDY 36

The Role of Investment in Two French Inflations. By H. LUBELL . 47

Industrial Relations in Nineteenth-Century Birmingham. By A. FOX 57

Retailing Costs: Some Comments on the Census of Distribution, 1950. By S. POLLARD and J. D. HUGHES. 71

Price Formation in Manufacturing Industry and Excess Capacity. By H. R. EDWARDS 94

OXFORD
AT THE CLARENDON PRESS
1955

OXFORD ECONOMIC PAPERS

(NEW SERIES)

OXFORD ECONOMIC PAPERS, of which eight numbers appeared at irregular intervals between 1938 and 1947, became a periodical in 1949 and is now published three times a year. The pre-paid subscription price is 25*s*., post free, and the price of a single copy 12*s*. net. Orders may be placed with any bookseller or sent direct to the Publisher.

PAPERS should be sent to the General Editor, c/o The Clarendon Press, Oxford. Contributors receive 50 free offprints. All correspondence other than that dealing with contributions should be addressed to the Publisher:

GEOFFREY CUMBERLEGE, OXFORD UNIVERSITY PRESS
AMEN HOUSE, LONDON, E.C. 4

AMERICAN ECONOMIC REVIEW

CONTENTS

VOLUME XLIV — DECEMBER 1954

INTERNATIONAL TRADE—THE SOVIET BLOC — Edward Ames
EXCHANGE RATE ADJUSTMENTS AND LIVING STANDARDS — W. L. Smith
OPTIMAL ADVERTISING AND OPTIMAL QUALITY — Robert Dorfman and P. O. Steiner
THE INTERNATIONAL SUGAR AGREEMENT OF 1953 — B. C. Swerling
PLANNING IN CONSUMER PURCHASES—DURABLE GOODS — Robert Ferber
INCOME SHARES: GREAT BRITAIN AND THE UNITED STATES — A. M. Cartter

REVIEW ARTICLES:
FULL EMPLOYMENT—1954 MODEL — P. J. Strayer
SCHUMPETER'S HISTORY OF ECONOMIC ANALYSIS — Jacob Viner

Reviews of Books, Titles of New Books, Periodicals, Notes

The AMERICAN ECONOMIC REVIEW, a quarterly, is the official publication of the American Economic Association and is sent to all members. The annual dues are $6.00. Address editorial communications to Dr. Bernard F. Haley, Editor, AMERICAN ECONOMIC REVIEW, Stanford University, Room 220, Stanford, California; for information concerning other publications and activities of the Association, communicate with the Secretary-Treasurer, Dr. James Washington Bell, American Economic Association, Northwestern University, Evanston, Illinois.

Send for information booklet

SOME IMPLICATIONS OF THE GROWTH OF INVESTMENT BY INSTITUTIONS, 1938–53

By RAYMOND FROST

THE purpose of this article is to re-examine Keynes's views, as found in the *General Theory*, on the behaviour of investors with regard to liquidity and to the valuation of shares and business assets, in the light of the growth of institutional investment since before the war. It will deal firstly with the statistical evidence of and the reasons for the increase in the amount of private savings represented by retained company profits, and by the increment in the investable funds of life assurance companies, pension funds, and building societies; secondly with the role of liquidity preference in the investment behaviour of these institutions; and finally, with the influence of their investment behaviour on the valuation of shares and business assets. It should be borne in mind that the comparison made in this article is between the behaviour of institutional investors as it is today, and that of individual investors as it was in the period about which Keynes was writing—the 1920's and 1930's. The earlier period was quite unlike the present, particularly in the market for shares. In the twenty years after 1913 the ownership of the bulk of British industry was converted from a family to a market basis by the floatation or issue of quoted industrial securities on an unprecedented scale; the nominal amount of quoted securities issued by home industry, as represented by the 'Commercial Industrial, &c.' category of the Stock Exchange official list increased three-fold between 1913 and 1933, i.e. from £439 million to £1,216 million.

This spate of issues took place in circumstances of eager demand and speculative excitement; the large increase in the supply of such securities did not prevent a doubling of the price level of equities between 1913 and 1939. Demand may have been too eager; it brought about a general over-valuation of equities and offered golden opportunities to the fraudulent company promoter. With the slump there came an equally excessive undervaluation of equities. It is not therefore surprising that Keynes, in the chapter on 'The State of Long Term Expectation', should have described the behaviour of investors in devastatingly caustic terms. What Keynes did not foresee was that with the growth of experience of this relatively new field of investment, the improvement of company law and the Stock Exchange regulations, the restriction of speculative credit, and the increase in investment costs, the behaviour of investors would be informed by a better and more sober judgement of industrial prospects. It is against the background of this improvement in the behaviour of

individual private investors that the growth of institutional investment should be assessed if one is not to ascribe to it a disproportionate influence amongst the various forces which have brought a new stability to the markets and to the economy over the last fifteen years.

The Growth of the Investment of Private Savings through Institutions

Keynes made no allowance in the *General Theory* for the respects in which the investment behaviour of institutions differs from that of individuals, possibly because the scale of institutional investment was insufficiently great for the differences to have any notable effect in the market at the time he was writing; today, however, such investment has grown to the point where its influence is distinct and recognizable. The principal institutional channels for the investment of personal savings are building societies, life assurance companies, and private pension funds other than those administered by life assurance companies. The growth in the amount of personal savings invested through these institutions is shown in the table below[1]

Net Increase (£ *million*)

Personal savings invested through	*1938*	*1945*	*1946*	*1947*	*1948*	*1949*	*1950*	*1951*	*1952*	*1953*	*Totals 1945–53*
1. Ordinary and industrial life assurance . .	60	93	111	113	123	126	133	134	149	187	..
2. Private pension funds.	20	30	30	40	50	60	75	90	100	110	..
3. Building societies .	43	26	49	72	78	98	110	97	118	157	..
Total of 1, 2, and 3 .	123	149	180	225	251	284	318	321	367	454	2,549

The retained profits of companies are the other main source of institutional finance; the growth in the proportion of retained to total profits is shown in the table below (Company Appropriation Account Table).

As the result of the growth in the proportion of profits retained, and also of the increase in savings by public corporations and authorities (which is, of course, another kind of institutional saving) the proportion of personal to total savings has declined since before the war (Table of Ratio of Personal to Total Savings).

[1] These figures are abstracted from Mr. C. T. Saunders's paper on 'The Pattern of Savings and Investment' read to the Manchester Statistical Society on 10 Nov. 1954, which he has very kindly allowed me to see before publication. The estimates for pension funds had not hitherto been published; they may turn out to be on the low side.

Company Appropriation Account. Percentage of Profits before Tax[1]

	1938	*1946*	*1947*	*1948*	*1949*	*1950*	*1951*	*1952*	*1953*
Distributions . . .	64·1	34·3	30·1	26·2	25·9	23·8	21·9	26·6	27·2
U.K. taxation . . .	7·9	29·3	31·2	32·2	30·5	29·3	35·7	32·5	29·7
Foreign remittances and taxes . . . / Retained profits after tax	28·0	36·4	38·7	41·6	43·6	46·8	42·4	40·9	43·1
Total profits (£ million) .	1,036	1,818	2,180	2,320	2,351	2,823	3,189	2,884	3,153

Ratio of Personal to Total Saving[2]

1938	*1946*	*1947*	*1948*	*1949*	*1950*	*1951*	*1952*	*1953*
%	%	%	%	%	%	%	%	%
54	54	8	3	5	4	9	31	34

The cause of the increase in savings by industrial companies is the breakdown of shareholdings, and hence of voting power, owing to the marketing of the greater part of industrial securities. This has increased the power of management to retain a larger proportion of profits whether shareholders like it or not.

The cause of the increase in savings through building societies and life assurance companies is the growth in the savings capacity of the lower and middle income groups who are accustomed to invest their savings in this way rather than directly through the market. This in turn reflects redistributive policies. It is possible to indicate with some precision the connexion between the distribution of wealth and the use of investment intermediaries. Mr. J. R. Cuthbertson has shown that in the periods 1909–13 and 1935–9 the proportion of insurance policies in deceased estates was over 10 per cent. in estates with a gross capital value of between £1,000 and £2,000, declining to under 2¼ per cent. for estates over £300,000.[3] A recent Harvard inquiry into a sample of American investors showed that only 10 per cent. of those with more than $500,000 of property held more than one-tenth of their assets in the form of life assurance, compared with 50 per cent. of those with assets of less than $25,000.[4] As for building societies, it is hardly necessary to labour the point that they are

[1] *Annual Abstract*, 1954. The 1953 figures are from table 293 (company appropriation account) which does not show figures for 1938; these have been taken from table 284 (corporate income appropriation account) which includes the income of public corporations —an item so small in 1938 that it can be legitimately ignored. Also, it is only since the war that 'Foreign remittances and taxes' have been shown as a separate item in table 293; they amount to less than 5 per cent. of total profits, and the proportion has if anything declined since 1938; I have not, therefore, kept them separate in the table.

[2] *Annual Abstract*, table 287.

[3] 'The Property Holding Habits of Individuals II', *The Banker*, Aug. 1953, p. 106.

[4] Butters, Thompson, & Bollinger, *Effects of Taxation: Investments by Individuals*, p. 325.

investment agencies for the small saver. The average of individual shareholdings in 1953 was only £511.

The increase in taxation as such has also made it advantageous for all income groups to invest through such media as building societies and life assurance companies, which have certain privileges in the way of tax relief.[1]

Liquidity Preference and the Characteristics of Investment by Institutions

In his discussion of the valuation of shares, and of the role of liquidity preference in the market for bonds, Keynes did not take account of certain special characteristics of investment by institutions, the existence of which I shall try to establish in this section.

It is necessary to anticipate at this point certain implications of the analysis which follows. Keynes evidently intended the theory of liquidity preference to apply exclusively to transactions involving *lending*; it is a theory of the determination of interest rates only in the sense of the rates of return on loans (rather than shares). Some dissatisfaction has been expressed about this restriction; Mr. T. Wilson, for example, refers to 'the main highway for speculative funds—in the more usual sense of the term—which runs directly between the short-term market and the market for industrials'.[2] Keynes may have assumed a division of investors between those who invest in bonds and those who invest in ordinary shares; on this basis it would be reasonable to provide an explanation for these groups in terms which were most appropriate to each. As little is known about the holdings of individual investors, this assumption seems as good as any. But a great deal is known about the holdings of financial institutions—in particular about those of the life assurance companies which are the largest group of all. And it is evident that they invest in both bonds and shares; and not only in these markets but also directly in industrial property. The distribution of their investments is dictated partly by considerations of safety, but within this restriction the tendency of their investment activities is to switch indifferently between different long-term investments till yields become equalized. In discussing their behaviour, it therefore makes little sense to confine references to liquidity preference exclusively to their activities in the bond market. And when dealing with the investment by industrial companies of their retained profits, it is

[1] The distributions paid by building societies to shareholders and depositors with a stake of less than £5,000 in any one society are taxed at a special composite rate which worked out last year at just over 30 per cent. Life assurance companies are liable to income-tax on profits at only 7*s.* 6*d.* in the £, and two-fifths of the amount of premium payments by the policy holder is deductible for the assessment of his liability to income and surtax. The Inland Revenue is in such cases adding to what the institutions can earn from their investments, thus increasing the benefits which they can offer to their clients.

[2] 'A Reconsideration of the Theory of Effective Demand', *Economica*, Nov. 1947.

evident that they do have motives to liquidity which are not primarily related to the bond market at all, but to uncertainty about the future yield of industrial assets—of factories, plant, and stocks. The role of liquidity preference in institutional investment will therefore be considered not only in connexion with transactions in bonds, but in relation to the full range of their investments.

I shall deal with only two groups of institutional investor—firstly, industrial companies, and secondly life assurance companies and pension funds. These two groups are quantitatively the most important channels of institutional investment in securities, and industrial assets. The approach to the detailed analysis of the investment behaviour of these two types of institution is determined by the distinguishing characteristics of their business. The business of an *industrial company* is to produce and trade; it therefore has to undertake trading risks. A trading loss implies a running down of working capital; and it is not easy to realize at their initial cost such assets as remain at the end of a period of loss. An industrial company is only able to bear such risks because a large proportion of its investment in assets is financed out of share capital and retained profits. The management of an industrial company is therefore principally concerned with current trading income and expenses; by the time it has to worry about capital repayment the game is up. The behaviour of an industrial company will therefore be analysed principally in terms of considerations of current income and current costs.

The business of a *life assurance company* is to borrow certain capital amounts from policy holders and invest them at a profit, repaying the principal plus the aggregate profit at a foreseeable date in the future. It is possible to eliminate a great deal of uncertainty from this business, and that is why a life assurance company can assume the heavy fixed obligations associated with what is in effect the debt finance obtained from policy holders. The largest elements in the transactions of an assurance company are capital ones: the raising of the principal of 'borrowings' (by way of premiums) and the realization of the principal of assets in order to repay them. Consequently more attention will be paid to the influence of capital transactions in the analysis of assurance company investment behaviour than in dealing with industrial companies.

Industrial Companies

Industrial companies have to take account of the uncertainty of future trading conditions in the investment of their resources, and in this respect they have motives to liquidity analogous to those which affect the private individual. But they lack the degree of opportunity which is open to the private individual for realizing a desire for liquidity. By private individual,

I mean the wealthy rentier whom Keynes appears to have had in mind. As his obligations are for the most part personal, he is legally able to reduce them; as his standard of living is high, it is possible for him to do so. He can in consequence afford to go liquid, and forgo income. The obligations of a company are far less flexible: the extent to which it can go liquid by reducing the scale of its activities and converting its operating assets into cash is limited by the relatively rigid structure of its current expenses and by the highly illiquid character of a large proportion of its assets. Normally its liquid position reflects passively changes in material prices and trading conditions which are outside its own control. The disposal of cash once it is acquired is, however, a matter of some relevance.

Let us assume that a company holds or expects to hold a certain amount of cash on some date, and that the management has decided not to invest it in trading assets. The alternative employments are to repay debt, to purchase securities, to pay it out in dividends, or to keep it in the bank. The table (below) gives some indication of decisions by the management of public companies as to the disposal of cash.

Aggregate Company Accounts

	I Mid-1951–mid-1952	*II Mid-1952–mid-1953*		
	Ratio to gross assets	*Ratio to gross assets*	*Change between periods I and II*	
Liquid Assets or Liabilities	%	%	£m.	%
(1) Marketable securities	3·8	3·7	− 18	− 4·0
(2) Cash	8·7	8·9	+ 75	+ 7·2
(3) Current liabilities (other than to bank)	20·6	19·5	− 14	− 0·5
(4) Bank overdrafts	3·8	3·2	− 54	−12·0
(5) Working capital*	31·9	31·8	+ 90	+ 0·2
Gross Assets (minus intangibl[illegible]	100·0	100·0	+603	+ 5·1
Ordinary dividends			+ 26	+12·3

* Total current assets — marketable securities, cash, together with stocks and debtors not shown separately above, *minus* current liabilities except bank overdrafts.

This table was prepared by Mr. N. Momtchiloff of the Industrial and Commercial Finance Corporation on the basis of the aggregate company accounts appearing every quarter in *The Economist*. The figures used in the table are from accounts published in 1952 and 1953, the only years in which it is possible to get a comparable sample. The very broad assumption has been made that the accounts are on average published six months after the completion of the companies' financial years. The figures for cash and marketable securities have been adjusted to eliminate the effect of the high proportion of the total of these items which is held by the trusts and finance houses in this sample of companies.

On this basis, the table shows that these assets rose by £169 million between the two financial years,[1] of which £75 million was held as additional cash, £68 million was used to repay debt, and £26 million to increase ordinary dividends. The discrepancy between the increase of £125 million in net liquid assets shown in the table, and the increase of only £90 million in working capital would represent the reduction of quoted securities by £18 million, and the reduction or writing down of stocks and debtors by £17 million. What seems to be the most remarkable feature of the table is that the liquidity ratios should be so stable, and the net changes in the amount of liquid assets so small over a period within which there occurred the final stages of the Korean war boom, the slump in the textile trade, and a reversal of monetary policy. If liquid assets showed such small changes over this period, they are likely to be no less stable in quieter times. Finally, it seems fair to suppose that the liquidity ratios in the table above are likely to overstate the true position, owing to the fact that physical assets are normally written in the books considerably below their cost or market values.

Life Assurance Companies and Pension Funds[2]

With life assurance companies and pension funds it is possible to make the quite definite generalization that by far the greater part of liabilities is fixed, and must be met. Share capital in the conventional sense is a negligible element here; the total issued share capital of the 86 ordinary life assurance companies established in Great Britain was at the end of 1952 only £39·3 million compared with a total life fund of £2,660·5 million. If one wishes to make an analogy with company finance, a life assurance company is principally financed by policy holders, to whom there is an obligation which cannot on any account be avoided. What the company is selling, and what the policy holder is buying, is financial certainty, and it would be inconceivable for the company to default upon that. The life offices' inescapable obligation is represented by the life fund, or the principal amount due at any date to policy holders. This, as we have seen, amounted to £2,660·5 million in 1952. The relatively flexible element in their capital structure is represented by the total of share capital, reserves and profit and loss accounts (which amounted in 1952 to £382·5 million) and above all, by the with-profit element in with-profit policies. In respect of the with-profit element on his policy (if he has such a policy) the policy holder is in the position of a shareholder; in respect of the principal sum assured, he is in the position of a bondholder. The bondholder element

[1] i.e. increase in cash and dividends plus *fall* in current liabilities and overdrafts.

[2] For the purposes of this article, the following analysis of life assurance company behaviour can be taken as being applicable to pension funds, which have, however, certain behaviour characteristics of their own.

predominates, and so in consequence does the proportion of 'safe' investments in the portfolios of the life offices. It is hardly surprising that one principle of life assurance investment is to endeavour to match the amount of the life fund with safe investments; the wonder is that they are willing to go into industrial investment to the extent that they do. The motive for riskier investment is to improve their earnings, and hence the benefits which they can pay on with-profit policies. This table shows the proportion of assurance companies' life funds covered by safe investments:[1]

1937	*1948*	*1949*	*1950*	*1951*	*1952*
%	%	%	%	%	%
102	105	106	108	107	107

But the disadvantage of a rigid structure of obligations is offset by the predictable way in which they mature. Actuarial calculations enable an assurance company to foresee each year within a small margin of error how much cash it will have to find at any date for many years ahead in order to meet maturing obligations. Given this information, it can then invest equivalent amounts in securities which are due to mature at or about the dates at which the obligations fall due. These obligations are of course compounded of an element of premium payments by policy holders and an element of compound interest earned by the investment of those premiums. The only other element of uncertainty is the future course of the interest rates at which future premium payments will have to be invested. This can be offset to a considerable extent by calculating the benefits on a policy issued today on the basis of reasonable and conservative expectations about the course of interest rates over the expected life of the policy. Thus an assurance company has no uncertainty about its future requirements of cash, beyond what may be required to cover the margin of error in actuarial calculations; and if it can find investments of appropriate maturities, it is not concerned with any uncertainty that may exist about the future course of the interest rates. It is not, therefore, affected by the two principal motives to liquidity (the 'precautionary' and the 'speculative'), except at times when interest rates are expected to increase in the very near future. Clearly the investment manager of a life assurance office would not buy a stock at 80 today if he thought he could get it at 75 within six weeks. But possibility of postponement is small, because of the company's remorseless need for income to meet obligations which have been calculated on the basis of a constant revenue from long-term investments. The life offices may switch their investable income from

[1] Calculated from Board of Trade Summaries of Assurance Business, balance sheets of companies established in Great Britain. The figure of 'safe' investments has been reached by deducting the total of investments in ordinary and preference shares from total assets.

one long-term market into another for a time (say, out of gilt-edged into industrial property) but they cannot afford to refrain for long from any kind of long-term investment. Another objection to delay is the sheer weight of their investable income. The investment manager of the Legal and General for example has to invest nearly £400,000 of net income a week; if he holds back for even a few weeks he will have an embarrassingly large quantity of money to dispose of, which may well turn prices in some markets against him if he tries to invest it quickly—and time is always pressing. The problem for each individual company is made more difficult by the general acceptance of price leadership. The number of life companies established in Great Britain is only 86; and while they do not consciously combine to influence the market, the majority tend to follow the leadership of the few giant companies in matters of investment. Thus the hands of the investment manager of the Prudential are tied not only by the great weight of the company's own funds, but also by those of the companies which will be following his decisions. For all these reasons the tendency towards steadiness and regularity in the investment of life assurance companies' income is very strong.[1] But if an assurance company is restricted in the extent to which it can refrain from investment out of income it is even more restricted in the extent to which it can switch out of investments into cash. Eighty-six companies control nearly £2,500 million of quoted investments, but it is obvious that the existence of a quotation for these investments does not imply that they can actually be realized *en bloc*.

Thus life assurance companies do not need to hold any considerable amount of cash, and would indeed find it inconvenient to do so. The table below shows the ratio of cash to total assets in the aggregate balance sheets of the life assurance companies established in Great Britain.[2]

1937	*1948*	*1949*	*1950*	*1951*	*1952*
%	%	%	%	%	%
2·0	3·0	2·8	2·6	2·8	2·9

To sum up the behaviour of institutions with regard to liquidity, it can be said that in the case of industrial companies the motives to liquidity may be present sometimes and to some extent, but the opportunity is restricted owing to the rigidity of their structure of assets and obligations; while with life assurance companies and pension funds the opportunity is

[1] It is for these reasons also that the institutions like to purchase new issues; it is always difficult to buy large amounts of existing securities without raising the price against the purchaser. A large proportion of the industrials purchased by institutions are either existing securities offered privately in large lots, or new issues, with the result that the institutions have come to dominate the new issue market.

[2] Calculated from Board of Trade Summaries of Assurance Business, balance sheets of companies established in Great Britain.

perhaps even more restricted, but the motive hardly exists at all. Bearing in mind the enormous weight of investable funds at the command of these institutions, this surely has some implications for Keynes's view about the role of liquidity preference in the economic system.

Let us now turn to the way in which the growth of savings through institutions implies some modification of Keynes's description of the way shares and industrial assets are valued. We have seen that assurance companies are not concerned with short-term movements in the market because they do not normally aim to realize their investments. It is not necessary to do more than compare this statement with the chapter on 'The State of Long-Term Expectation' to see that in so far as the share market is influenced by the activities of the assurance companies and pension funds, that chapter is no longer an adequate description of market behaviour.

The life offices increased their holdings of ordinary shares by £50 million in 1953,[1] and it may well be that the pension funds are investing as much again in this way. By comparison, total ordinary share issues in 1953 amounted to £103 million.[2] Direct purchases of industrial property are also being made by pension funds, and to a lesser extent by life assurance companies; this type of investment does, of course, involve the direct valuation of a business asset, and it is undertaken by institutions on lines which bear small resemblance to the Keynesian picture in the chapter mentioned above. And the principal source of finance for industrial development is the retained profits of industry and this implies that the most influential set of estimates about the prospective yield of business assets is made by people who have to take an irrevocable long-term decision.

[1] Statement by Chairman of British Insurance Association, 25 June 1954.
[2] *The Economist* figures for new capital issues.

THE REHABILITATION OF MONETARY THEORY IN SWEDEN

By ERIN E. FLEETWOOD

PROFESSOR ERIK LUNDBERG'S book *Business Cycles and Economic Policy*[1] is very hard to review. It is a large and most ambitious work which deals with a great many aspects of the central theme. Moreover, Professor Lundberg's main points are often so carefully qualified that it is sometimes difficult to know exactly what the author's opinion really is. This caution is probably partly due to the fact that he knew that in Sweden his book would be revolutionary; and it has, in fact, already given rise to a widespread discussion on economic policy which will probably have far-reaching effects. Indeed, the development of economic theory in Sweden is bound to be affected as a result of the new slant given to the discussion by Professor Lundberg and not least by his revival of older theories in new forms. The book reflects the author's struggle to arrive at new conceptions, as far as Swedish economic thought is concerned, and to work out a fresh method of approach, but it cannot be said to constitute, as now presented, a clear statement of fully matured ideas. Nor can it be said that many of the ideas expressed by Lundberg are 'new' in an absolute sense, as most of them are already fairly familiar to economists in England and on the Continent. In Sweden, however, where, for instance, the interest rate was banned as a subject of discussion for several years, from the spring of 1945 onwards, many of Lundberg's conceptions are new to most of those who have taken up the study of economics since the late 1920's. Only an economist of Erik Lundberg's standing could give these old ideas in their new form sufficient authority to establish a new trend; and he is to be congratulated on having had the courage to show so fully in what direction his economic thinking has been taking him.

The Stockholm School of Economics had its origin in the lectures and writings of Knut Wicksell, a man of the Left with, on the whole, liberal ideas, and in the work and teaching of David Davidson,[2] politically a

[1] *Konjunkturer och ekonomisk Politik*, Stockholm, 1953 (550 pages).

[2] Knut Wicksell's first major published work was *Geldzins und Güterpreise* (Jena, 1898). This was followed by the *Lectures on Political Economy*, which were published in Swedish in 1901 and 1906 and appeared in an English translation in 1934, 9 years after Wicksell's death. David Davidson reviewed Wicksell's *Geldzins und Güterpreise* in *Ekonomisk Tidskrift* in 1899 and in this review he already made clear the main points on which he differed from Wicksell. Although there was a vast field of agreement between Wicksell and Davidson, the debate continued for over two decades, mainly in articles in *Ekonomisk Tidskrift*. Gustav Cassel and Eli F. Heckscher also contributed to the economic theory of their time, but they did not occupy quite such a central position as Wicksell and Davidson, who, in spite of all their differences, had the most profound respect for each other's economic thinking. Wicksell had the more original mind, and it is probably true to say that he is now recognized all over the world as having been one of the half-dozen greatest economists of the last hundred years.

Conservative with a great veneration for Ricardo, but free of all dogmatism, as is shown by his analytical *exposés* and his own contributions to economic thought. In the late 1920's and early 1930's a new generation of economists—somewhat anxious not to be overwhelmed by their great predecessors—won fame for their macro-economic models, including the ex-ante and ex-post type of analysis of total savings, investment, demand, &c., and for their break-downs of these aggregates. There were, of course, certain divergencies in the views of these younger economists, but it may be said that, generally speaking, the School tended to become increasingly sceptical of the value and possibilities of monetary and credit policy both under conditions of depression and under those of full boom; on the other hand, great stress was laid on the application of fiscal (i.e. budgetary) and physical measures of control for the direction of the economy along the lines desired by the political powers. The political aims to be pursued included the maintenance of a high and steady level of employment and also, as far as possible, the attainment of a more even distribution of income. The general attitude of this second generation no doubt influenced Sweden's economic policy in the post-war period; but the theoretical and practical studies which were made during the war and post-war periods began once again to develop along new lines, some economists being influenced more particularly by Schumpeter and some adopting an econometric approach. The Stockholm School as such seemed suddenly to have lost its vitality, one reason being that several of its leading members had become heavily engaged in political work, so that they had little time left for other activities. Professor Lundberg's book may be regarded as a most serious attempt to revive the Stockholm School in a modified form by reincorporating into its body of thought certain of the ideas of the founders, with special stress on the importance of monetary policy and thus also of savings, and by adapting the quantity theory to modern economic conditions and to the 'political possibilities' of the 1950's. The strength of this *exposé* lies in the fact that it is so largely built on the author's own practical experience gained when acting as consultant to war and post-war administrators; that is, however, also its weakness. For, with the best will in the world, Lundberg's book cannot be said to present a coherent set of principles, based on a continuous chain of reasoning. It is a contribution to current thought—not really a finished product.

It would be quite impossible to review within the limited space of a short article all the problems dealt with by Lundberg in this book. Attention must therefore be concentrated on some of his main ideas, and these will be divided for this purpose into three main groups: firstly, his view on the cyclical fluctuations during the past 35 years, which he gives more or less as an introduction; secondly, his analysis of the problem of the

inflationary gap and, in connexion therewith, his assessment of the usefulness of measures of fiscal and monetary policy; and, thirdly, his estimate of the value and efficacy of controls in relation to the price structure.

Fluctuations in Business Activity during the Period 1918–53

Lundberg's analysis of the international business cycle between the two world wars runs along familiar lines and is, of course, chiefly concerned with conditions in the United States; but the development in that country is seen against the background of numerous factors making for instability in the world economy, especially during the 1920's.

In Lundberg's opinion 'it seems very probable that the American crisis and depression was chiefly initiated and determined by factors having their origin in the domestic business trend in the United States; it must, however, be admitted that the international business decline and not least the financial crisis of 1931 acted as "aggravating circumstances" '[1] (p. 25). In this connexion Lundberg seems to have overlooked the fact that from around 1925 onwards the U.S. economy was directly and indirectly affected by the depression in the overseas raw-material-producing countries, and especially in the food-producing countries. As far as his own explanation is concerned, moreover, he cannot be said to have made it sufficiently clear that the productive capacity of the United States had been left unimpaired by the war and that there had been an amazing investment boom which was bound to come to an end eventually because of the diminishing real return of all the large-scale investments which had been undertaken. The United States was thus 'ripe' for an industrial depression, and the coincidence of this state of affairs with a world-wide agricultural depression, a building slump in the United States, and a financial crisis both at home and abroad certainly accentuated the downward trend.

With regard to the international aspects of the question, Lundberg thinks that the depression in the United States was the cause of the world depression. In so far as export-import relationships are of decisive importance, and in so far as the depression psychosis tends to spread from nation to nation, the decline in the United States undoubtedly influenced the trend in other countries. But Lundberg has nowhere clearly indicated that, unlike the United States, the European countries were not 'ripe' for a depression of their own. On the contrary, he finds that 'in marked contrast to the development after the Second World War, the business cycles in the various countries were highly *synchronized* during the interwar period' (p. 36). This categoric statement not only conflicts with information given by Lundberg in other parts of his book (e.g. p. 87), but it also shows that Lundberg has lost sight of the important facts that by 1929 the backlog of

[1] This and all subsequent quotations from Lundberg's book have been translated by the reviewer.

demand in the countries of Western Europe resulting from the war had still not been satisfied, that in these countries the economic structure was reasonably stable and that internal demand was well maintained, as was evident even after 1929. The multiplier effect of steady domestic demand for consumer goods—and, in some countries, even for durable consumer goods—sustained the European economies, which had not been subject to 'overinvestment'. In fact, building activity revealed its oft-noted anti-cyclical tendency to increase in periods of reduced costs. Sweden, though very dependent on international trade, had in 1930 and some of the following years a large volume of internal demand—which, however, as in Great Britain, caused difficulties in the balance of payments.

Lundberg's analysis of developments in Sweden itself starts with a survey of what might be termed long-term factors. He shows how in Sweden, productivity—that uncertain term which here is interpreted as real national income per person in the labour force—although lower after the First World War than before it, has been consistently higher than in most other countries. As regards the reasons given for the relatively high level of productivity in Sweden, there is a very interesting section (p. 55) in which the author discusses the increase in total productivity which took place in connexion with the movement of labour from the land to industry, the building trade, and other branches of production. While in 1910 45·2 per cent. of Sweden's actively employed population was engaged in agriculture, in 1949 the proportion was only 21·8 per cent. Those workers who left agriculture could expect to earn, on an average, twice as much in their new occupations. The statistics which Lundberg gives are only intended to serve as an example, but as such they are interesting and help to bring out a point which is very often overlooked in similar analyses. The whole of the inter-war period was characterized in Sweden by an overall increase and also a *per capita* increase in the *real* national income, which expanded at a remarkably steady rate throughout the twenty years from 1919 to 1939, the increase corresponding not only to an improvement in real wages, but also to the persistent movement of workers from the less well-paid occupations (mainly agriculture) to the better paid (industry and services).

The depression of the 1930's therefore assumed a comparatively mild form in Sweden. It was superimposed on a persistent upward trend; and it is, consequently, a point particularly worthy of note that in the case of Sweden the depression was imported from abroad, chiefly because of a sudden and substantial decrease in the value and even the volume of the country's exports. The volume of industrial production in Sweden decreased by only 13 per cent. between 1930 and 1932, while the unemployment figure reached its peak of some 200,000 in January 1933. Internal consumption remained reasonably stable, though there was a slight decline

in the demand for durable consumer goods. The persistent upward trend in the Swedish economy has previously been remarked upon by Lundberg,[1] and his findings have now received independent confirmation from a study by E. M. Phelps Brown and S. V. Hopkins[2] in which it is found, with regard to wage-rates, that 'between 1923 and 1939, in Sweden, the United Kingdom and the United States of America, they were rising each year on the average by something like 1·7 per cent. of their mean value in the latter 1920's' and also that over long periods 'in every country except Sweden (in 1860–1914) there is a higher proportion of rises in the real wage-rate in the phase of rising activity; but the contrast between the rising and the falling phases is not sharp, for real wage-rates also rose in more than half the years of falling activity'.[3] In Sweden real wages rose as often in years of stable (intermediate) or falling business activity as in those characterized by an upward trend.

It was exports which, by the expansion in their volume in 1933, came to the rescue of the domestic business trend 'about half a year before it was possible to speak of definite signs of a general upswing in Swedish economic activity' (p. 77). By reason of their buoyancy, exports have again and again given the first impetus to a revival of business in Sweden; this is partly due, as Lundberg points out, to the favourable composition of the country's exports, for which there is a very wide market. Though Lundberg is fully aware of the important part played by exports in helping to overcome the depression in Sweden and points out that Great Britain was the chief market for Swedish goods, he does not seem to have laid nearly sufficient stress on the great advantage gained by Sweden from the devaluation of the krona in 1931, especially in view of the fact that this devaluation put an end to the fall in prices and thus furnished a basis for the restoration of balance in the internal economy and in relation to other countries. If the Swedish krona had not been devalued, Sweden would presumably have been exposed to very much the same difficulties as were experienced by the gold-bloc countries. Had the krona not been gradually linked to the pound, Sweden would not have been able to reap all the benefits which it did from the building boom in Great Britain, nor would it have derived so much advantage as it did from the general recovery in the sterling area, which took place at a time when the dollar area was still in a state of depression which lasted, with only brief intervals of revival, right up to the outbreak of the Second World War.[4]

[1] *Det Svenska Näringlivets Konjunkturkänslighet* (Meddelanden från Konjunkturinstitutet, serie B: 5), Stockholm, 1945, p. 8.

[2] 'Wage-Rates in Five Countries 1860–1939', *Oxford Economic Papers*, 1950, p. 236.

[3] Ibid., p. 246.

[4] The undervaluation of the krona in the 1930's, which some reviewers of Professor Lundberg's book have made much of, did in fact exist and probably helped somewhat to facilitate Sweden's recovery. It was not, however, a planned measure deliberately adopted,

It is interesting to note that in Lundberg's opinion the *immediate* effects of the much publicized expansion which began in 1932, when the Socialists took over the government in Sweden, were of relatively minor importance. The volume of state investment did not begin to increase appreciably until 1934–5 (one reason for the delay being a prolonged strike in the previous year) and by that time the upturn of the business cycle was already past its first stage. Lundberg himself says that 'the increase in state investment activity which was actually brought about must have exerted a relatively slight influence as a stimulus to the business trend' and he adds that 'of much greater importance for the upturn of the business cycle, which clearly started in the summer of 1933, was the expansive combination of foreign exchange and monetary policies, this combination being a prerequisite for the rapid recovery of the private economy' (p. 107). Notwithstanding the stress laid again and again on the true causes of the Swedish recovery—earlier by Marquis W. Childs[1] for instance, and now by Erik Lundberg—Sweden is still often cited as an example of a country which carried out with success an anti-depression policy of public works in the 1930's.

The Inflationary Gap: Fiscal and Monetary Policies

The theoretical discussions of the problem of the inflationary gap and of the importance of fiscal and monetary policies have not been concentrated in a few chapters but are to be found in a number of different places throughout the book. Lundberg introduces a theoretical analysis whenever, in examining the measures which have been taken, he feels the need to bring out their implications from the point of view of theory. For instance, the multiplier is analysed in some detail in connexion with the description of the business cycle in Sweden, which is thus found to fit in with Lundberg's model; this method results in a clarity of exposition such as is seldom encountered. Equations and mathematical explanations are mostly relegated to footnotes, as are, unfortunately, also some very valuable comments and a number of well-chosen quotations from other authors.

One disadvantage of this method is, however, that it leads to a certain amount of repetition and also—what is more serious—that it gives the impression that the various aspects of economic policy discussed in different parts of the book are considered against the background of the same set of conditions, whereas in fact this is not always so. To mention but one example: whereas in one place in his book Lundberg deals with the views of Fisher, Cassel, and Keynes on the quantity theory, in examining which Keynes, in particular, stressed not only the consequences of changes

but one of the results of the confusion caused by the Kreuger crisis, which followed Ivar Kreuger's death in Paris in March 1932.

[1] *Sweden, the Middle Way*, Yale, 1936, p. 146.

in interest rates but also the direct effects of limitations on the availability of new credit (pp. 306 ff.), in the chapter on wages (pp. 397 ff.) the possible development of wages is discussed in terms of inflationary conditions, and not on the assumption that a non-inflationary policy is being pursued, i.e. a policy providing for the limitation of new credit granted.

Lundberg examines in an extremely interesting manner the usefulness of the concept of the 'inflationary gap' and of the drawing up of a 'national budget' as guides to economic policy (pp. 321, 372). In this field he is outstanding, having been concerned with these problems over many years as head of the Swedish Institute for the Study of the Business Trend. To go adequately into what he has to say on the subject is not possible here, but his conclusions may be briefly summarized. He emphasizes that the theoretical problems involved in this type of approach are far from being fully solved as yet; and when the nature of the statistical information is examined it is found to be even less reliable than the theory would lead one to expect—which helps to explain the numerous incorrect ex-ante prognoses. Lundberg, therefore, warns most definitely against the indiscriminate use of the method of inflationary-gap analysis as a basis for the shaping of economic policy. The inherent uncertainty is too great. This sceptical attitude is to be found throughout his book; it colours his positive approach in discussing the role of monetary and financial policy in solving economic problems, as well as his negative attitude towards controls as tools of economic policy.

When discussing the importance of financial policy for the general balance of the economy, Lundberg points out that it is not sufficient simply to compare the budget figures with the gross national income but that a more detailed analysis is required. With regard to the degree of liquidity and the money volume,[1] however, he finds that it is the budget figure on both income and capital accounts which is decisive; it is the relationship between the cash position and income and expenditure which will determine the degree of liquidity not only of banks and other financial institutions but also of individuals. Compared with previous Swedish thinking on the subject this attitude is almost revolutionary.

In his analysis of the influence which the budget exerts on the economy, Lundberg points out that the state's demand for goods and services, compared with the total demand for goods and services, and also the state's share in total gross investments are much better measures of the size of the budget and give a truer indication of its effect than a simple comparison with the gross national income. In his calculations he admits, however, that that part of budget expenditure which represents transfers of

[1] See below for the importance attached by Lundberg to the quantity theory and its implications with regard to economic and especially monetary policy.

income from one income group to another (or from one sector of the economy to another), although it does not affect the state's demand for goods and services, does in fact tend to increase total demand, especially when the transfers are made to low-income groups.

Similar arguments are used to show that the existence of a budget surplus or deficit, though it is of importance to the Finance Minister from a political point of view, does not necessarily have the income-destroying or income-creating effect which it is generally assumed to have. The statistical tables which are given to illustrate this point are impressive. Unfortunately, in the case of the table which shows the ex-post position as regards the effects of the state budget on the formation of private income in the years 1945–6 to 1948–9, it is not clear whether full allowance has been made for the changes in the cash position, or whether the figures relate only to the formal budget position. As there were considerable accumulated appropriations in the accounts of the various departments, the evolution of the 'cash position' is even more important in Sweden than in many other countries; it may in fact deviate considerably from the formal (even the ex-post) results of the state budget accounts.

All this uncertainty is further increased when account is taken of the complicated effects of the multiplier. Lundberg discusses multiplier reactions at various stages and ends by pointing out that 'our knowledge of the reaction mechanisms of the economy is so incomplete that attempts to make mathematical "precisions" must give a very rapidly diminishing return'. 'There is, therefore,' he adds, 'in my opinion no reason to favour financial policy in preference to monetary policy on the grounds that "you know much better what you are doing"' (p. 387). This is also the general conclusion reached by Lundberg in his discussion of the practicability of making use of differential rates of taxation or a varying scale of subsidies in the case of, say, exports and imports and housing, or in order to affect conditions for certain commodities, for instance by according different treatment to margarine and butter. He emphasizes that too little is known about the price mechanism and how it functions for the administrator to be able to tell exactly what will be the effects of any given tax or subsidy, as it is often not even known whether the market price of the commodity in question is too high or too low. Yet if the price function is interfered with the reason for doing so must be to attain a clearly formulated aim and the repercussions of the measures envisaged must be known in advance. But these requirements are almost impossible to fulfil; and *it is therefore advisable that in using financial policy as an instrument of business-cycle policy only methods which are very general in their effects should be employed, e.g. a general turnover tax, general investment taxes, &c., as opposed to specific, individual measures affecting special sectors of the economy.*

In his analysis of monetary policy Lundberg falls back on the quantity theory, discusses its origins, its development in Sweden by Wicksell and Cassel, and also its history in England, where Keynes's treatment of the theory in *A Tract on Monetary Reform* (1923) and *A Treatise on Money* (1930) is judged to have contributed more to its development than the attitude which he adopted in *The General Theory*.

The impreciseness of the conception of velocity of circulation (so difficult to measure statistically) is, of course, brought out again as the weakest element in the quantity theory. But his conclusion is that the time has probably come to create a synthesis of (i) the quantity theory (in its modernized form); (ii) Keynes's theory concerning the correlation between the availability of cash and the interest rate; and (iii) Wicksell's theory concerning the effect of the interest rate on the volume of investment. Whether one considers private persons, business firms, and banks individually or takes society as a whole, there is, of course, undoubtedly a connexion between the availability of cash and the propensity to spend. In view of the 'strategic' position which this connexion occupies in economic policy 'the *control of the money volume*, i.e. of the total money supply, *should be a means whereby the central bank can affect the society's expenditure and income and thereby indirectly also the level of prices and the volume of employment*' (p. 308; reviewer's italics). It is because Lundberg has adopted this line of thinking that he is so critical of the attitude taken by the Sveriges Riksbank after 1945 that it was 'powerless' to influence the monetary affairs of the country by monetary means. He even goes so far as to call the Riksbank interventions in the post-war period a manifestation of 'bad conscience', since again and again it urged the government to stiffen direct (physical) controls, thus, in his view, diverting attention from its own ineffective monetary policy.

Though Lundberg perhaps errs on the side of caution in his statement as to how effective monetary policy can be in the various phases of the business cycle, he is convinced that, given favourable conditions, relatively minor monetary measures are capable of influencing the behaviour of banks and firms. He thinks that this is especially the case if there exists *some degree of uncertainty as to the methods* which will be used by the Central Bank to give effect to its policy; but the aims of this policy, e.g. the maintenance of a constant price level, should, he thinks, be stated as clearly as possible. Such a statement concerning the aim pursued is considered necessary for the purpose of producing the psychological effect, the creation of which is one of the objects of monetary policy. It must be added, however, that Lundberg concentrates his discussion rather one-sidedly on the conditions obtaining in a boom. Maybe he regards it as a foregone conclusion that in our days the tendency will be all in favour of

a cheap-money policy, so that the question of what monetary policy should be adopted in a depression needs no special examination.

As far as *practical* policy is concerned, Lundberg rates the importance of monetary policy very high, considering it to be one of the main general means of controlling economic development. The reasons why he so clearly prefers general monetary policy to 'functional finance'—i.e. detailed and more or less differentiated tax measures—are:

firstly, because he has come to realize that monetary policy is *not* ineffective; that changes in the money supply do matter and that small changes in interest rates and *general* (as opposed to detailed) measures of credit control are effective, influencing the whole economy, and not only the few specific points which are all that can be reached by measures of physical control (unless the physical measures in question are so comprehensive that they amount to total control);

and secondly, because his analysis of the primary effects and, much more, the secondary effects of specific taxes and subsidies has shown him that these measures interfere with the working of the cost and price structure, a mechanism so delicate that it cannot safely be tampered with.

Controls and the Price Structure

The task Lundberg originally accepted when the work which led to his present book began was to study the effect of controls over imports, prices, and building activity, with special attention to the effect on industry. Indeed, the starting-point of *Business Cycles and Economic Policy* was an inquiry into what industrial and commercial firms thought of the existing controls and how they reacted to them in practice. Information was collected by means of questionnaires, supplemented by personal interviews, and the results have been summarized in an appendix. The exact value of such an inquiry is, as Lundberg points out, difficult to assess, since there is a limit to the number of firms which will actively assist; and the firms which do take part often seize the opportunity—especially in interviews—to voice their complaints. The inquiry has, however, thrown some new light on how controls really function and has thus helped to make the study realistic. Some of the results obtained are used in two chapters which deal, in particular, with the subject of controls; in one of these chapters the author describes how controls function in the various branches of the economy and in the other he discusses the way in which they affect the price structure. His detailed study of the way in which trade, investment, and price controls have been administered leads him to the conclusion that, whereas in war-time abnormal conditions may make possible the exercise of detailed control, under peace-time conditions the apparatus of

control, even when it is functioning at its best, is not skilful enough, nor sufficiently quick and flexible, to ensure efficiency of administration. The lack of flexibility mentioned above makes it difficult, of course, to adjust controls to changing circumstances, and the rules of thumb which generally have to be used do not permit the often very important secondary effects of the controls to be taken sufficiently into consideration. Consequently a system of detailed controls offers no possibility of making allowance for distortions of the economic structure and provides no standard which can serve as a guide for an optimum utilization of the factors of production.

Lundberg explains why this is so in a subsection entitled 'The Dilemma of the Economy of Controls' (pp. 484–5), from which the following may be quoted:

> The essential reason why the economic benefits to society as a whole of the interventions made by means of controls could not be clearly demonstrated in the examples given above is to be found in the impossibility of surveying from a central point the effect of the interventions which, in the last instance, are reflected in the market reactions of firms and consumers. Since—even in a controlled economy—these market reactions are the result of the weighing-up of considerations of prices, costs and profits, the determination of prices is the decisive indicator of the social economic effect. And, as the various examples have shown, *the dilemma inherent in the system of controls lies precisely in the fact that the price mechanism has been largely put out of operation as a regulator of economic activity at the very time when an attempt is being made to evaluate the social-economic results of controls with the help of a study of price movements.* (Lundberg's italics.)

In other words, when prices no longer reflect the market conditions of supply and demand they can no longer be used as yardsticks for estimating the national income or forming an opinion regarding changes in the trend of consumption and production.[1] In this connexion Lundberg points out that the most eminent theoreticians among socialist economists (e.g. Oscar Lange and Abba P. Lerner) have themselves built their systems around effectively functioning price mechanisms corresponding to those operating in a competitive market economy.

The free formation of prices has, according to Lundberg, important dynamic functions to fulfil in that it will encourage 'increased production or imports in certain places and cut down activity in others' (p. 489), and also because it allows a continuous process of building up profits and indicates where the possibilities of profit lie. Were price control so efficient as to eliminate all 'abnormal' profits there would be no change in the pattern of production (and investment) in the direction indicated by the change in demand.

[1] Lundberg's reasoning is not merely a theoretical construction; it has a very practical application. It may in this connexion be of interest to mention that the trade treaties currently concluded between countries within the Soviet bloc are reported to use for their price arrangements the prices quoted on the free markets of the Western world.

It is precisely in order to permit the free formation of prices that Lundberg advocates *general* measures of control, i.e. measures of monetary and financial policy, instead of direct (physical) controls, functional finance, and detailed credit control. Specific (as opposed to general) measures of control require in order to be successful:

firstly an economic policy with well thought-out and clearly defined aims;

secondly, a reliable yardstick to guide the choice between alternative uses of resources; and

thirdly, great accuracy in forecasting economic developments.

These, however, are conditions which are not fulfilled today and probably never will be. Above all, the essential yardstick could be provided only by prices formed in competitive markets, such a guide being a necessity for any rational control system built on centralized economic planning. But, as stated above, *the dilemma* is that the effectiveness of the process of price formation is undermined by the control system itself, which is specifically designed to prevent unpredictable market reactions. And Lundberg furthermore thinks that if it is assumed that such a state of affairs exists that the three conditions indicated above are fulfilled, then 'one can with good reason question the need for applying a strongly interventionist economic policy in order to attain general objectives relating to the business trend' (p. 494).

This conclusion is reinforced by Lundberg's analysis of the probable working of controls during a depression. Whilst on the one hand the authorities responsible for the price control would be expected to reduce prices, on the other hand they might be tempted to stabilize certain prices at a relatively high level, at the expense of the consumers, in order to protect certain important branches of the economy—for instance, agriculture. Whether such a price stabilization would result, at the same time, in a stabilization of employment, as Lundberg implies, is, however, a moot point. American experience in the New Deal period of the 1930's shows that the rate of employment was actually highest in those industries in which price codes were non-existent or were not observed.

Moreover, as Lundberg points out, the control of investments is really only control of building activity which affects only some 25 per cent. of the volume of gross investment. On the other hand, activity in the other branches of investment, especially investment in inventories and industrial machinery, is likely to undergo considerably larger fluctuations than building, and thus be a much more serious threat to economic stability. As for import control, though it can be useful in overcoming a temporary foreign exchange crisis, there is always a tendency for it to be continued even after the immediate crisis is over; indeed, failure to correct the situa-

tion by fiscal and monetary measures (including movable interest rates) is always likely to prolong the original disequilibrium.

If monetary policy were tightened up, for instance, the general conditions for investment would become less favourable and investment expenditure in different spheres would be lower than would otherwise have been the case; th effect of the reduction in investment on the flow of income would, furthermore, bring about a fall in expenditure on consumer goods and in the volume of imports.[1] Lundberg is at pains to point out that the particular features of this development 'depend on a multitude of adjustment processes, unanalysable in detail, in the numerous markets where demand and supply meet' (p. 472). Physical controls, on the other hand, cut right through the market system and represent a form of dictatorship which seeks to take the place of the adjustment which comes about in the framework of relatively free market conditions.

Though Lundberg stresses that physical controls, detailed control of credit, and functional finance disturb the working of the markets and shatter the price structure, he should perhaps have held on to this line of reasoning sufficiently strongly to be able to avoid resorting to a series of clearly political arguments in his final summing-up. Be this as it may, it is difficult to quarrel with him when, quoting Adam Smith, he emphasizes that the desirability of a society in which each individual will have the least possible opportunity for creating mischief, especially the numerous mediocre and very fallible politicians, administrators and economists whom we shall always have with us. In Lundberg's view, these requirements are best fulfilled by an economy relying largely on the operation of decentralized markets, and possessing a price mechanism which is allowed to function reasonably freely, such an economic system thus being a safeguard against the danger of too great a concentration of power in any one hand. Evidently he feels obliged to resort to these political arguments, for no sooner has he stated the case for allowing the price mechanism to function freely than he proceeds to point out the obvious weaknesses of such a system in our times, the most outstanding among which is the existence of cartels and other kinds of restrictive practices. Hence, while he is certainly moving towards an increased appreciation of the advantages of a price system built on free and effective markets, his approach is very hesitant, presumably because he entertains some doubts as to whether such a price system can be made a reality.

It is typical that Lundberg should end his book with the statement of

[1] This is the *immediate* effect; but the reviewer would like to draw attention to the fact that in those countries in which monetary confidence has been maintained by a flexible interest-rate policy and in other ways the flow of domestic savings has generally become more abundant, thus making possible a larger volume of investment than could otherwise have been achieved.

an economic dilemma. One of the signs of his intellectual indecisiveness is that there is no concluding chapter, but merely some 'Final Observations' included as part of the last chapter. While this attitude—it is only fair to add—may reflect his capacity for seeing simultaneously both the good and the bad points in almost all theories and policies, it seems a pity that he should have such difficulty in recognizing the need—even for an economist—to try to arrive, at some stage, at a definite judgement if any effective policy is to be pursued. It is this lack of a definite opinion that leaves the reader finally with a feeling of vagueness, but it does not, however, prevent *Business Cycles and Economic Policy* from being an important contribution to the body of economic literature. It brings together in one volume an enormous amount of theory and facts, and the treatment of many of the particular points investigated will open new vistas and stimulate further inquiry. While the author's endeavour to provide Sweden with a new approach to economic thinking leaves an impression of incompleteness, it is clear that his considerable experience and his insight into the workings of controls in all their forms, on the one hand, and his political predilection for a free society, on the other hand, have caused him, on balance, to favour the rehabilitation of monetary policy —and have at the same time led him to re-examine many of the basic concepts and relationships of monetary theory.

BASLE CENTRE FOR ECONOMIC AND FINANCIAL RESEARCH.

THE INTRANSITIVITY OF CERTAIN CRITERIA USED IN WELFARE ECONOMICS

By W. M. GORMAN

SUMMARY

In this paper it is shown that neither the Scitovsky nor the double index number criterion for an increase in welfare is transitive. The consistent use of either could therefore lead to inconsistent judgements. The relationship of these results to Arrow's general theorem concerning the impossibility of constructing a social welfare function satisfying certain restrictions is discussed. The author feels sure that many of these results must be familiar to individual economists, but they do not appear to be generally known.

1. Introduction

1.1. The purpose of this paper is to clear up a few technical points in welfare economics, not to present yet another critique of the subject. Nevertheless a few brief introductory remarks may put things in better perspective.

1.2. The three basic ideas in Welfare Economics are that social choices should be based on private preferences, that one situation should be chosen rather than another if each individual prefers it, and that the economist, as such, has no right to compare changes in the welfare of different individuals. That these ideas are rather muddled has been ably shown by others; here we merely note that they are rarely sufficient to permit the economist to recommend one action rather than another.

1.3. Kaldor first attempted to evade this difficulty by suggesting that a proposed action should be taken if the potential gainers could compensate the potential losers and still have something over. Such a criterion clearly favours change—for it compares the actual situation ruling before the change, not with the actual situation which it would bring about, but with the most favourable of a set of potential situations. Indeed Scitovsky has shown that an economist who based his judgements on it might find himself recommending first a new policy then its reversal. He therefore suggested his double criterion—that a proposed policy should be adopted if it satisfied the Kaldor criterion, while its later reversal would not.

1.4. The basic concept in all this is that of 'compensation'. As Mr. Bailey has recently discussed it in the *Economic Journal* it is not necessary to enter into details here, but a few remarks may be of use.

1.5. In the first place it seems clear that we accept the Kaldor criterion as intuitively reasonable because we think of compensation as a costless process, and feel that, if the gainers by a change are not made to compensate the losers, when they could do so, it must be because society prefers it so.

Suppose we call social preferences 'right-minded' if they depend only on individual preferences, and if situations preferred by every individual are preferred by society. Then, viewed in this light, the Kaldor and Scitovsky criteria may be regarded as tests for the consistency and 'right-mindedness' of society's preferences. Figure 2 illustrates the proposition that they are inadequate tests—for it is easily seen that the Scitovsky criterion would imply that b should be preferred to a and a' to b'—though people are individually precisely as well off in a as in a', and in b as in b'. It is therefore to be expected that, though the Scitovsky criterion will guard the welfare economist from the crude contradictions mentioned in paragraph 2, it may itself lead him into similar contradictions in more complicated cases. That Scitovsky realized that this might be so seems clear from the following quotation.

> It is true that as we have seen such a change would improve general welfare for some welfare distributions and worsen it for others; on the other hand, we are not interested in all possible welfare distributions. *There are only two distributions of welfare which really matter.*[1] Those actually obtaining immediately before and after the change contemplated. It seems therefore sufficient to concentrate on these and to investigate how the change would affect general welfare if it were to leave the distribution of welfare unaffected, and if that were both before and after it, first what it actually is before, secondly what it actually is after, the change. Whenever these two comparisons yield identical results, we can make welfare propositions of almost the same generality and significance as the allocative welfare propositions of the classical economists; especially since the identical results for the two welfare distributions imply a strong presumption in favour of the same result holding for all intermediate welfare distributions as well.

It would appear that Scitovsky did not visualize the routine use of his criterion. It is only in a single application that 'only two distributions of welfare really matter'.

1.6. It is usual to assume that any physically possible distribution of welfare can be attained by redistribution. Broadly, this implies that compensation be paid by means of lump sum taxes and bounties and that there should be perfect competition. This assumption will be needed for some of the results in the present paper. In general, however, it will be sufficient to associate with each situation $\alpha, \beta, \gamma, \ldots$, not only the actual distribution of welfare $A, B, C, \ldots$, but also a set of distributions to be thought of as potentially attainable from it by 'redistribution'. The upper bounds of these sets will be called the corresponding 'utility possibility loci' $a, b, c, \ldots$. These will frequently be determined by physical and technical conditions. It is clear that it would be meaningless to apply the compensation criteria to the comparison of suboptimal distributions—that is, where everyone could be made better off by mere 'redistribution'. Moreover, the results

[1] Quoted from reference III. My italics.

proved below are trivial in such cases. We will therefore confine our attention to optimal distributions, so that each distribution will lie on the corresponding 'utility possibility locus'.

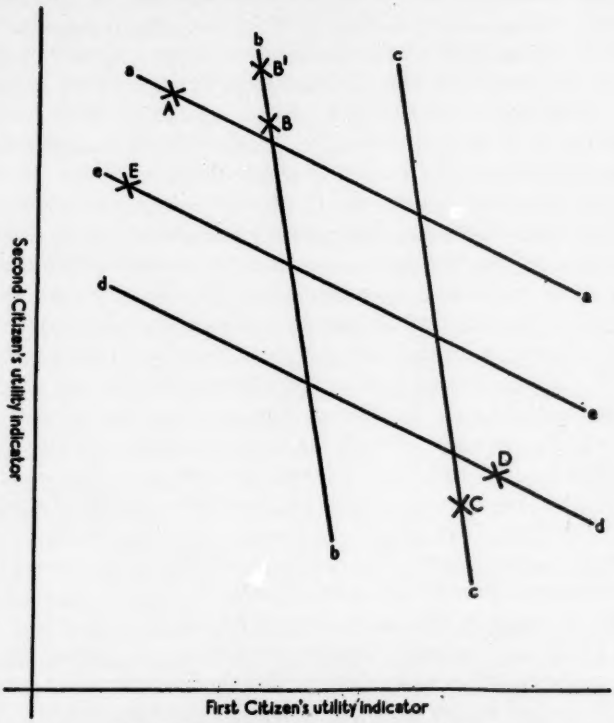


FIG. 1. A portion of a utility possibility map

2. The Real Compensation Criteria

2.1. Throughout this section we will refer to two-dimensional utility possibility maps for illustration. This is done solely to simplify exposition. The arguments can easily be generalized to cover the general case.

2.2. Consider now the situation illustrated in Figure 1. Clearly by redistributing income at B we can move to B', where everyone is better off than at A. On the other hand it is impossible to redistribute income at A in such a manner as to make everyone better off than they would be at B. Hence, according to the Scitovsky criterion B is to be preferred to A. It will be convenient to write this $A < B$ and say that A is inferior to B. Similar arguments show that $B < C$ and that $C < D$. Were the criterion transitive we would therefore have $A < D$. A direct comparison of A and

D shows that $D < A$. This contradicts the previous result and shows that the criterion is not transitive.

2.3. To understand the implications of these results we imagine a government which bases its policies on the Scitovsky criterion. At the outset, we will suppose, the economy is in situation A. A new policy is suggested which would lead to a move to B. The Scitovsky criterion is applied. It is seen that $A < B$, and the new policy is put into effect. Later the possibility of moving to C becomes apparent. The change is made, for the same reasons as before. Next D is offered and chosen. Finally a policy is suggested which would lead to a move to E. Since $D < E$ the policy is approved. By a series of free choices we have moved from A to E. Yet, not only is everyone worse off in E than in A, but the utility possibility locus lies uniformly below that at A. That is, however the government may redistribute welfare in the final position, there would be a redistribution of the original welfare in which everyone would be better off. Obviously its economic advisers have involved it in difficulties very similar to those already discussed in the introduction.

2.4. It is of some interest to investigate the conditions under which these 'contradictions' are possible. We will take a given utility possibility map and ask whether the applications of the Kaldor criterion, or of the Scitovsky double criterion, to positions on this map can lead to a 'contradiction'. It will be shown that such contradictions are always possible if any pair of utility possibility loci cut. That this is true of the Kaldor criterion is obvious. To see that the same is true of the Scitovsky criterion under reasonable continuity assumptions, let the intersecting loci be the curves a and b in Figure 1. Let then c be a slightly displaced version of b in which slightly more of each good is available.[1] Similarly, slightly less of each good is available along e than a, and slightly less again along d. The configuration so constructed is obviously similar to that of Figure 1. By choosing the points A, B, C, D, and E in similar positions we can derive the same 'contradictions' as before. Hence, given reasonable continuity assumptions, we can always construct situations in which the Scitovsky criterion would yield bad 'contradictions'. Now this is just the case in which the Kaldor criterion leads to 'contradictions'. In an obvious sense the two criteria break down together; the Scitovsky criterion is never sufficient to guarantee the 'validity' of a decision unless the Kaldor criterion alone is sufficient to do so.

3. The Index Number Criteria

3.1. The Kaldor and Scitovsky criteria can only be applied when the utility possibility maps are known, which is rather a tall order. The closely

[1] If the utility possibility loci refer to situations in which the total supply of each good is not given, but rather the productive techniques and quantities of factors, we let slightly more of each factor be available in c than b. Similar changes should be made below. Similar arguments may obviously be applied in more general cases.

related index number criteria, on the other hand, are based on data which is more likely to be available.[1]

3.2. The first index number criterion states that b is better than a if:

$$\sum p_b q_a < \sum p_b q_b, \cdots, \tag{i}$$

that is, if the total money expenditure in situation b would have been sufficient to purchase the goods actually bought in situation a. Here p_a, p_b are the prices ruling in a and b, and q_a, q_b the amounts actually purchased at these prices. Unfortunately, we might have $\sum p_a q_b < \sum p_a q_a$ as well as (i)—that is, the first index number criterion might contradict itself in the same manner as the Kaldor criterion did. In order to guard against this possibility it may be required that a second index criterion also should be satisfied, namely that:

$$\sum p_a q_a < \sum p_a q_b \cdots . \tag{ii}$$

3.3. The economic justification for these criteria has been discussed elsewhere—notably by Samuelson (reference IV). The discussion will not be repeated here. However, two points must be noted. Firstly, it is clear that their use can only be justified in the comparison of positions of competitive equilibrium—i.e. Pareto optima—for it is only there that prices have a simple welfare meaning. Secondly, they are more obviously useful in the comparison of exchange equilibria—where total supplies are given—than in that of general equilibria—where a change in the pattern of production is possible. For the present, therefore, we will confine our attention to cases of exchange equilibrium.

3.4. The index number criteria are evidently closely related to the compensation criteria of Kaldor and Scitovsky. To bring out this relationship we will write $a <^* b$ whenever the double index criterion is satisfied—i.e. (i) and (ii) both hold. To compare the criteria in more detail we must have recourse to the community indifference map[2]—since prices and price lines do not appear on the utility possibility map. The convexity of the community indifference maps is easily shown to imply that the Scitovsky and double index number criteria cannot contradict each other—i.e. it is

[1] Strictly these criteria are used in evaluating a change which has already taken place. It is only in such cases that the data is available.

[2] *Notation:* We have already seen that each 'situation' has to be represented by an actual distribution of welfare and by a utility possibility locus. In the case we are at present discussing the utility possibility locus is determined by the supplies available, while the community indifference curves are determined by the actual distributions of welfare. For this reason we have denoted the latter curves by capital letters A, B, C,..., and the points representing the supplies available by the corresponding lower-case letters a, b, c,... . In this notation it is easy to see that b is to be preferred to a on the Kaldor criterion if b lies above the community indifference curve A. Similarly $a < b$—on the Scitovsky criterion—if B lies above a but A does not lie above b. The first index number criterion is satisfied if a lies below the price line through b, while $a <^* b$ implies also that b lies above the price line through a.

impossible to have $a < b$ and $b <^* a$ both holding at the same time. On the other hand, Figure 2 illustrates the proposition that neither is just a weaker version of the other. Here we see that $a < b$, because b lies above A, while

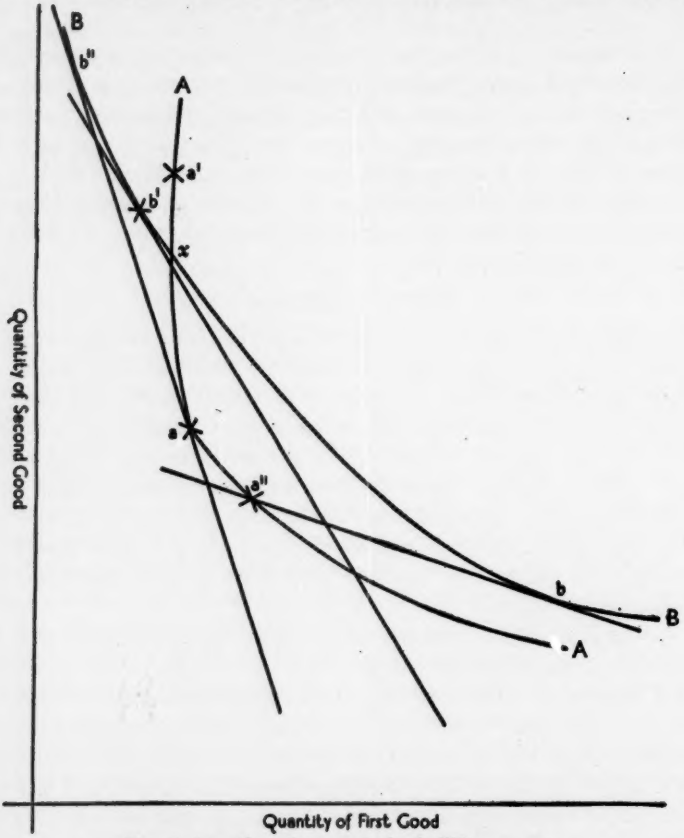


FIG. 2. A portion of a community indifference map

a does not lie above B. Yet $a \not<^* b$ since both a and b lie above the price line through the other. On the other hand, $a <^* b'$, since a lies below the price line through b' while b' lies above that through a; while $a \not< b'$ since a lies above B, and b' lies above A.[1]

[1] It is reasonable to ask whether Figure 2 represents a possible configuration. We will indicate how it may be constructed given any pair of community indifference curves A, B intersecting at a point x. First we choose b', $b''a'$ close to x as shown and draw the tangent $b''a$ to A. Choose now a'' slightly below a as shown and draw the tangent $a''b$ to B. That this construction is always feasible can easily be shown given the usual assumptions that the indifference curves are convex and the marginal rates of substitution positive everywhere.

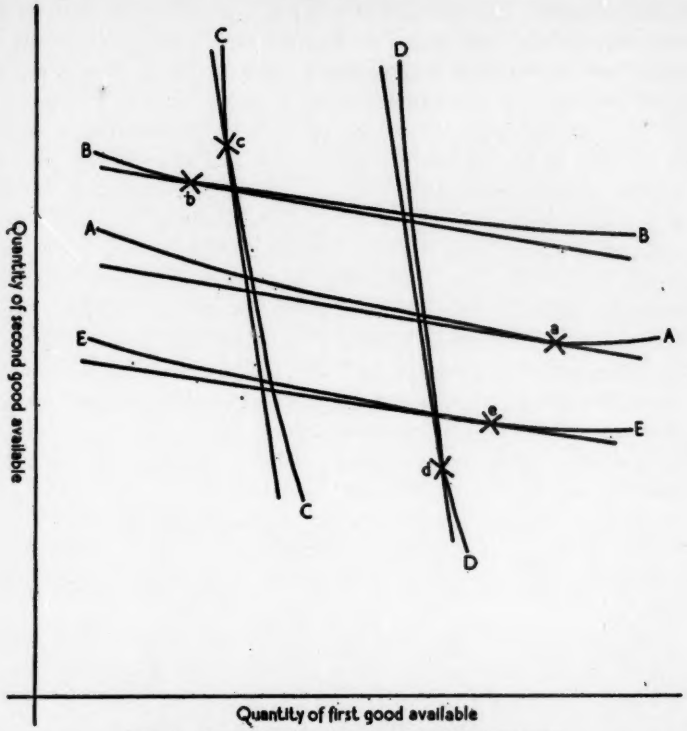


FIG. 3. A portion of a community indifference map

3.5. We will now prove the intransitivity of the double index number criterion in the case of exchange equilibrium by arguments similar to those we used to prove the intransitivity of the Scitovsky criterion. To do so, let us consider Figure 3, which represents a portion of the community indifference map for the economy in question. To keep the treatment simple, we have assumed that there are only two commodities, but the method of discussion is perfectly general. The index number criterion clearly yields $b <^* c$, since all the components of q_b are less than the corresponding components of q_a. Similarly $d <^* a$. Moreover, the price line at d passes above c, while that at c passes below d. Hence $c <^* d$. Similarly $a <^* b$. We now have $a <^* b$, $b <^* c$, $c <^* d$, and $d <^* a$. If the criterion were transitive the first three of these would imply $a <^* d$, which contradicts the fourth. Hence it is intransitive.

3.6. We will now show, on reasonable continuity assumptions, that, if any pair of community indifference curves cut, we can construct a situation in which the consistent use of the double index number criterion leads to

inconsistent results. The method of proof will be similar to that used in the case of the Scitovsky criterion. Let the intersecting community indifference curves be C and A in Figure 3. Let the curve B be a slightly displaced version of A in which everyone is slightly better off than in A, and D be a slightly displaced version of C, in which everyone is slightly better off than in C. Choose the points a, b, c, d close to the points of intersection of the curves, as in Figure 3. Clearly this can always be done. It is now easily shown at $a <^* b, b <^* c, c <^* d, d <^* a$, as before.[1] Hence the application of the index number criterion would lead the government to prefer b to a, c to b, d to c, and a to d, which is inconsistent. A stronger inconsistency can be constructed by drawing the curve E corresponding to a situation in which everyone is slightly worse off than in A. Clearly this can be drawn so that d is below it and e can be chosen on it as in the figure. We then have $d <^* e$. Hence the government would even prefer e to d, and therefore to a—though everyone is worse off in e than in a, and the utility possibility locus corresponding to e lies uniformly below that at a. A consistent application of the index number criteria would lead to a 'contradiction'. Clearly therefore we can only apply this criterion without fear of error if no pair of community indifference curves cut, in which case the problem is rather trivial. In this case the first criterion would suffice by itself.

3.7. We have already seen that each of the two compensation criteria break down in just those cases in which a pair of utility possibility loci cut. In the case of exchange equilibrium this will happen in just those cases in which a pair of community indifference curves cut. This is precisely the case in which the index number criteria break down.[2] That is, under exchange equilibrium the Kaldor, Scitovsky, first index number, and double index number criteria all break down if any one of them does so. Moreover this will happen if, and only if, a pair of community indifference curves cut—that is, if the individual citizen's Engel curves at any given set of prices are parallel straight lines.[3]

3.8. So far, discussion in this section has been limited to cases of exchange equilibrium. Clearly the methods are formally applicable to any situation in which price and quantity data are available. We have already seen that they are obviously of no particular significance unless the economy is in equilibrium—since prices have otherwise no welfare significance. There remains the case of general competitive equilibrium, where changes in

[1] The reason of this is that the indifference curves A, B, C, D (and E) approximate to straight lines—the price lines through a, b, c, d (and e). Moreover, those at c and d are approximately parallel, as are those at a, b (and e). The calculus proof is simple. The results can be extended to cover cases in which there are kinks in the indifference curves.

[2] i.e. the first criterion and the double criterion.

[3] Reference VI.

production are possible. The significance of the index number criteria in such situations has been discussed elsewhere. We will merely discuss the formal problem of transitivity. There are difficulties in doing even this, for general equilibrium is most easily discussed on the utility possibility map, where prices and price lines do not appear. It can, however, be shown that if production techniques are potentially continuously variable, then a necessary and significant condition for the index number criterion to break down is that a pair of potential utility possibility loci cut. This in turn is equivalent to the assumption that the utility possibility loci at some potential production optima cut, which is in turn equivalent to the cutting of a pair of community indifference curves in the region of potential production optima. Once again the compensation and the index number criteria break down in the came cases.

4. Social Welfare Functions based on the Criteria

4.1. The relationship of the above results to Arrow's theorem[1] on the impossibility of constructing a social ordering satisfying certain criteria, is of some interest.

4.2. In the first place we must note that none of the criteria discussed above yield a complete ordering of all situations. This is true for two reasons. In the first place it is usual to restrict comparisons to Pareto optima relative to given external conditions. This is imperative in the case of the index number criterion—which derives its value entirely from the assumption that individuals have reached their equilibrium position relative to the prevailing price. It is highly convenient in the case of the Kaldor and Scitovsky criteria—for, if a point is sub-optimal, everyone can be made better off by a mere redistribution. This clearly makes the Kaldor criterion completely useless; it could imply that we should move from one sub-optimal position to another in which everyone was worse off—not by a series of steps, but in a single one. The Scitovsky criterion is not quite so badly affected. However, the idea underlying it is surely that redistribution is a painless operation, and, if so, we should clearly first make such redistributions as make everyone better off—i.e. move to a Pareto optimum.

4.3. However, the criteria do not provide a complete ordering of even the Pareto optima. For instance, there will be many cases in which the Scitovsky criterion gives neither $A < B$ nor $B < A$. For a complete ordering it would be necessary to decide what to do in such cases. The simplest solution is to say that the two situations are equivalent when the criterion fails to show that either is preferred to the other. It is easily shown that such a complete ordering regarded as an ordering of all cases, optimal and sub-optimal, is intransitive in the sense that A may be equivalent to B

[1] Reference V.

and B to C though $A < C$. This is proved in Arrow's book (reference V, page 44). This is, of course, a weaker result than that proved in section 2.

4.4. Let us now consider the case where no pair of utility possibility loci cut, and only Pareto optima are to be discussed. To simplify the discussion assume that the supply of each good in each situation is given. In this case there exists a community indifference map which satisfies all the conditions usually laid down for personal indifference maps, and further, yields straight Engel lines. Applied to such a map we see that the Scitovsky ordering requires that we prefer to be on a higher, rather than on a lower, indifference curve, and are indifferent between situations on the same curve. It is clear that this is a genuine social ordering—it is transitive, and permits us to compare any pair of admissible situations. Moreover, it depends only on the personal orderings of the various citizens. By Arrow's general theorem on such orderings, it must fail to satisfy some of his conditions. In fact it satisfies all but one of them[1]—that the ordering should be independent of 'irrelevant' alternatives. This condition implies that the social ordering of any set of alternatives should depend only upon the manner in which the alternatives in question are ordered by the citizens, and be independent of their ordering of other alternatives. Now the Scitovsky criterion depends on the idea of *compensation*—which implies a knowledge of the indifference curves on which the various alternatives lie—that is, the ordering of certain further alternatives. Hence it violates this criterion. It might be argued that we are discussing a rather degenerate special case in which the Scitovsky criterion is equivalent to choice using a well-behaved indifference map. Viewed in this light, the difficulty is that the community indifference map depends on a knowledge of the *complete* personal indifference maps.

4.5. We might attempt to change the double index number criterion into a social welfare function in the same way as we did with the Scitovsky criterion, that is, by saying $a =^* b$, if neither $a <^* b$ nor $b <^* a$. It is then easy to construct cases in which $a =^* b$, $b =^* c$ while $a <^* c$ except where every single indifference curve of every individual is a straight line and all the straight lines are parallel. This social ordering is of little interest!

UNIVERSITY OF BIRMINGHAM

[1] This is not strictly correct—for two of Arrow's conditions are too strong for his purposes. Thus, for instance, they cannot be applied to cases in which the personal indifference curves are known to be continuous. A weaker form is strong enough to prove his results. In this weaker form, only one of them fails to be satisfied by the Scitovsky ordering above.

REFERENCES

I. N. KALDOR, 'Welfare Propositions in Economics', *Economic Journal*, September 1939.

II. N. KALDOR, 'A Note on Tariffs and the Terms of Trade', *Economica*, November 1940.

III. T. DE SCITOVSKY, 'A Note on Welfare Propositions in Economics', *Review of Economic Studies*, vol. ix (1941–2).
IV. P. A. SAMUELSON, 'The Evaluation of Social Income', *Oxford Economic Papers*, February 1950.
V. K. J. ARROW, *Social Choice and Individual Values*, Wiley, New York, 1951.
VI. W. M. GORMAN, 'Community Preference Fields', *Econometrica*, January 1953.

THE VALUATION OF NET INVESTMENT

By CHARLES KENNEDY

1. In 1941, just at a time when the nation's capital must have been depreciating at an unprecedented rate, there was carried on in the pages of *Economica* a fascinating debate between Professors Pigou and Hayek on the theoretical problems involved in the notion of maintaining capital intact.[1] In the following year Professor Hicks joined in,[2] with the declared purpose of provoking the two contestants to another round, and because in his view the question of maintaining capital intact just could not be left where they had left it the summer before. The challenge was not taken up, and Professor Hicks's short article remains, as far as I know, the last word on the subject.[3] It is now perhaps time for us to ask in our turn whether the question of maintaining capital intact ought to be left where he left it over twelve years ago. This article suggests that the answer is 'no', and attempts to grapple once again with this very difficult problem.

In this task I am conscious of a handicap, in that I am unfortunately incapable of understanding what is the positive contribution of one of the contestants, namely Professor Hayek, though I think I have grasped some of his critical arguments. I shall have therefore to substitute for his analysis an analysis of my own, though one which I think has some features in common with his. In so doing I may be doing him a serious injustice; but my sense of guilt is tempered by a comfortable feeling that I am sinning in company. Professor Pigou already stands accused of having misunderstood Professor Hayek's position in one not unimportant respect;[4] and I am not altogether sure that Professor Hicks is correct in interpreting him as wishing to define net income in terms of a 'constant income stream'.[5] Professor Hayek is clearly a difficult writer.

I propose to begin at the wrong end and to discuss Professor Hicks's solution of the problem first, because this will bring us most quickly to

[1] Pigou, 'Maintaining Capital Intact'; Hayek, 'Maintaining Capital Intact: A Reply'; August 1941.

[2] 'Maintaining Capital Intact: a Further Suggestion', *Economica*, May 1942; henceforward referred to in footnotes as '1942 article'.

[3] See, however, G. F. Break, 'Capital Maintenance and the Concept of Income', *Journal of Political Economy*, February 1954, for a very useful, but mainly classificatory, treatment of the subject.

[4] Hayek, op. cit., p. 279; Hicks, op. cit., p. 175 n.

[5] At least, I do not myself get this impression from Chapter XXV of Hayek's *Pure Theory of Capital* (Macmillan, 1941) to which its author refers us in his later article (p. 279). But then, to be candid, I cannot discover either the 'alternative procedure' for defining net income that we are there promised. Hayek's standpoint is to me very much clearer in an earlier article, 'The Maintenance of Capital', *Economica*, August 1935. There he does indeed use the 'constant income stream' hypothesis, but not without recognizing its arbitrariness (pp. 249–50).

fundamentals. Superficially, it would appear that his main contribution to the debate was the suggestion, imported from Sweden, that a distinction must be drawn between the depreciation or obsolescence in any period which was foreseen and that which was not foreseen. Thus, he argues,[1] to measure depreciation by subtracting from the value of the capital stock at the beginning of the year (C_0) the value of these same goods at the end of the year (C_1) is internally inconsistent, because the actual events of the year influence the figure for C_1 but cannot influence the figure for C_0 unless they have been correctly foreseen. Instead, we must replace C_0 by the value (C_0') which would have been put upon the initial stock at the beginning of the year if the events of the year had been correctly foreseen, including among those events the capital value C_1 at the end of the year. Depreciation is then equal to $C_0'-C_1$ and net investment to

$$\text{gross investment}+C_1-C_0'.$$

Now this distinction between foreseen and unforeseen changes is no doubt entirely reasonable and indeed for some purposes very important. But we must not allow this to blind us to the fact that the Hicksian solution rests also on a procedure which is very much more questionable, namely that of using the change in the market value of the capital stock between the beginning and the end of the year in the calculation of depreciation, no matter whether the former value is corrected in the manner suggested above or not. Indeed, as I hope to show, Professor Hicks's import from Sweden is, despite its fine wrappings, unsound at the core, and leads to some very curious results. We may get some insight into why this is so by taking an example.

Suppose that at the beginning of a year there existed a certain number of machines for the production of some commodity; and that these machines were expected to earn certain quasi-rents and therefore had a certain capital value. Suppose further that during the course of the year a similar number of machines of exactly the same type are built. Now, it is by no means theoretically inconceivable—however unlikely it may be in practice—that the competition of the new machines reduces the value of the existing machines to less than half their original value; and this remains a theoretical possibility even when the introduction of the new machines was foreseen at the beginning of the year. Yet, if we were to adopt Professor Hicks's suggestion, we should have to say that capital had not been maintained intact, even though in physical terms it had doubled! Such a contradiction surely presents a very serious objection to Professor Hicks's method, more serious indeed than that raised by

[1] Op. cit., p. 177.

Professor Hayek to Professor Pigou's device of measuring capital in physical terms.

The reason why we obtained so curious a result, the reason in fact why we must reject any use of the change in market values for measuring depreciation, is rather fundamental, not only to the problem in hand but also to the theory of value in general. It is of course none other than the familiar proposition that in economics values are determined at the margin. Whatever may be its other uses, the aggregate value of a flow of consumption goods has no claim to be regarded as a measure of the total utility of those goods. For the same reason, the aggregate value of a stock of capital goods cannot be regarded as constituting a measure of the total usefulness of those goods. Still less are we justified in using the difference between the aggregate values at two dates as a measure of depreciation.

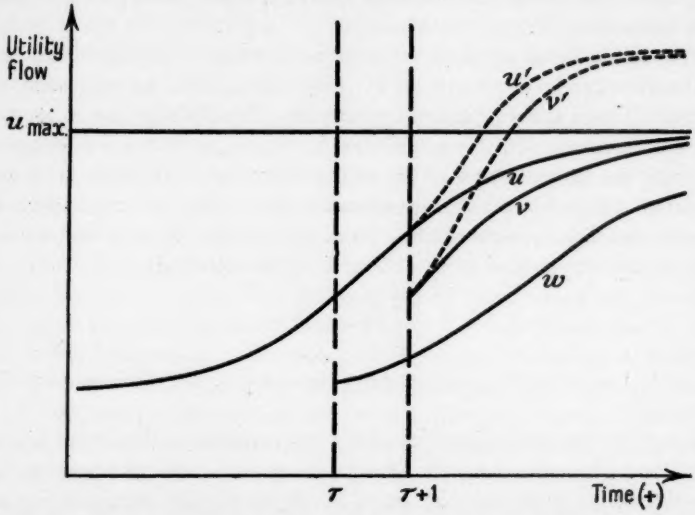
The failure of the Swedish definition when applied to aggregates presents us with a dilemma, and there are a number of courses we might follow. On the one hand, we can step boldly inside the margin and measure values in subjective terms. On the other hand, we may follow the two paths marked out by Professor Hicks in his celebrated 1940 article,[1] where aggregate money values are used, but where they take on an entirely different significance from that implied by the Swedish definition of depreciation. We shall explore each of these roads in turn.

2. The subjective road takes us at once over some very treacherous ground. We shall have to make many assumptions, including some which will be frankly unacceptable to most people and others which constitute very sweeping simplification. Thus, at the outset, we shall have to assume that utility is not only measurable but can be summed both for many individuals and over time. The reader who rightly objects to such an assumption may take comfort from the fact that this particular line of analysis fails in the end to lead to any satisfactory conclusion. Nevertheless, the reasons for its failure are, I think, instructive; and it is because of this that I must ask the ordinalist to bear with the assumption for a few pages. Further, we shall assume that we are dealing with a community made up of a constant number of individuals who live for ever; that there is no time preference, i.e. that utility in a future period is given the same weight as utility in the current period; that all decisions are in the hands of a benevolent dictator with perfect foresight, whose one aim is to maximize the total utility flow over time accruing to his subjects; lastly, that there are no changes in tastes or in productive techniques, even though some of the techniques have not at the outset been exploited to the full.

[1] 'The Valuation of the Social Income', *Economica*, May 1940; henceforward referred to as '1940 article'.

The assumption of constant techniques, and with it that of perfect foresight, will have to be relaxed later.

In passing, it may be as well to discuss very briefly the definition of net income by constant income (utility) stream, which Professor Hicks first put forward in *Value and Capital*,[1] but which he repudiates in his 1942 article. Quite apart from the objections which he raises, there is another reason why it is not of any help to us in our present purpose. For the definition of net income as 'the largest utility flow in any period which can be



permanently maintained in future periods' is, in an important sense, a *hypothetical* definition. It can only be used as a measure of net income provided the community is in fact consuming or expecting to consume all its income.[2] But, in practice, our dictator will want to do better than maintain the largest possible constant utility flow: he will want to increase the utility flow in the future by investing now. And, if net investment takes place, the definition by constant utility flow gives us no guidance as to how to measure either net investment or net income. For this we need another construction altogether.

Let us suppose that at time T our community has a certain stock of capital which is less than the optimum. One of the tasks the dictator has to perform is to determine the rate at which additions to the capital stock

[1] O.U.P., 1939. In Chapter XIV, where Hicks introduces the definition, he uses a money measure of net income and not a utility measure as we do here (for reasons which should appear later).

[2] For a further discussion and slight qualification of this criticism, see a later footnote.

should be made until the state of bliss is reached where the community is able to sit back and enjoy its maximum utility flow (u_{max}). Let us suppose further that he has solved this problem so that the course of future investment is already mapped out. This means also that the future course of the utility flow is laid down at the same time. At this point a diagram will be useful. Measuring the utility flow up the vertical axis and time (t) along the horizontal axis, we can draw a curve (u) denoting the course of the utility flow from time T onwards.[1] At the same time we may find it useful to project this curve backwards so as to trace the history of the utility flow before time T.

First of all, let us try to find a measure, in terms of utility, of the value of the total capital stock at time T. This is easily done by supposing that at time T there is no capital stock and that the dictator has to start, as it were, from scratch. He would then have to trace out a new curve (w) showing the future course of the utility flow; but, so long as we assume that our dictator had made no mistakes up to time T, the w curve will merely be a displacement of the u curve to the right. We can then measure the value of the capital stock at time T by the integral

$$\int_{T}^{\infty} (u-w)\,dt$$

or, graphically, by the area between the u and w curves from time T to eternity.[2]

Secondly, we need a measure of the net investment that takes place in the first period after time T. This is also easily found by supposing that the community finds itself at time $T+1$ in exactly the same position as it was at time T. The dictator would then have to trace out a new curve (v), which is once again a displacement of the u curve to the right, this time by one period only. We can then measure net investment in the period from T to $T+1$ by the integral

$$\int_{T+1}^{\infty} (u-v)\,dt$$

or, graphically, by the area between the u and v curves from time $T+1$ to eternity. It will be noticed that, since the v curve is a mere displacement of the u curve, this last integral is also equal to the integral

$$\int_{T}^{T+1} (u_{max}-u)\,dt$$

[1] We have drawn the u curve so as to approach u_{max} asymptotically; but the exact shape of the u curve is of little importance to the main argument.

[2] This integral will be finite provided (*a*) there is a definite limit to the utility flow at u_{max}, (*b*) it takes a finite time to reach the position at time T from a position where there is no capital.

or to the area between the u_{max} line and the u curve from time T to time $T+1$.[1] We then also obtain a simple measure of net income, defined as the addition of net investment and current utility. The net income flow during the period is in fact exactly equal to u_{max}.[2] Another reassuring thing about the measures we have fashioned, although this is conditional on the restrictive assumptions made, is that net investment during the period is equal to the change in the value of the total capital stock between the beginning and the end of the period, as may easily be seen by displacing the w curve to the right by one period.

Our device of measuring in subjective (utility) terms has enabled us to steer clear of the first hazard, the rock on which the Swedish definition at once foundered. This is of course the fact that, even though there is no change in the techniques of production, nevertheless any additions to the capital stock will bring about some 'obsolescence' (in terms of market value) in the initial capital. We have now to see whether we can take account of the other kind of obsolescence, that brought about by actual changes in technique. In the analysis of this problem, we cannot expect to obtain quite so pretty results as we did earlier, but we do nevertheless obtain results of a kind. But, before going on to discuss them, we must alter some of our assumptions: we now drop the assumption that production techniques remain unchanged, and we can no longer assume that our dictator has perfect foresight, though we shall still have to assume that at any one time his expectations are certain, even if not necessarily correct.

Suppose that just before time $T+1$ some invention takes place which entirely alters the prospects for the community, increasing its maximum utility flow in the state of bliss. The dictator will now have to alter his ideas and trace out a new curve (u') showing the future course of the utility flow. But we can still obtain an unambiguous measure of the net investment which has actually taken place in a period. To do so, all we

[1] As is readily seen by adding to the total utility flow between T and $T+1$ first the area between the u and v curves in the following period, then that in the next period, and so on.

[2] Some points of comparison between this measure and the 'constant utility stream' measure may be appropriate here. Both measures are, in the sense used above, *hypothetical* measures, the hypothesis being in each case the future course of the utility stream. But, for most purposes, it seems to me more sensible to take as hypothesis a future course that is aimed at or expected rather than one that in general nobody would aim at or expect. This is not to suggest that the 'constant utility stream' hypothesis is without significance in its own right, nor to deny that in ordinary life we use the word 'income' in this connexion—witness the expression 'living within one's income'. What I find misleading in Chapter XIV of *Value and Capital*, however, is the suggestion that only the 'constant utility stream' hypothesis can provide a measure of *ex-ante* net income. (Notice that, so long as we assume perfect foresight, our own measure of net income can be interpreted in either an *ex-ante* or an *ex-post* sense.) On the other hand, I should admit that it is the only hypothesis that does not involve us in the summation, as distinct from the ordinal comparison, of utility flows over time—and this is, no doubt, why Professor Hicks gave it the attention he did. Cf. also Hayek, 'The Maintenance of Capital', *Economica*, August 1935, loc. cit.

have to suppose is that at time $T+1$ the community finds itself in exactly the same position as it was at time T, except that the invention is now known to have taken place. The dictator would then have to draw a new curve (v') showing the future course of the utility flow, it being noted that this time the v' curve is not a simple displacement of the u curve or of the u' curve or of a combination of the two. We can now measure the value of net investment by the integral

$$\int_{T+1}^{\infty} (u'-v')\,dt$$

or, graphically, by the area between the u' and v' curves from time $T+1$ to eternity. This integral, which is clearly an *ex-post* measure of net investment, since it measures the gain in expected future utility from time $T+1$ onwards achieved *as a result of* the investment in the preceding period, is to be distinguished from

$$\int_{T+1}^{\infty} (u-v)\,dt$$

which still remains an *ex-ante* measure of net investment, since it measures the gain *expected* at time T to be achieved as a result of the investment planned. (In the earlier case where expectations remained unchanged, the *ex-post* measure coincided with the *ex-ante* measure.)

Certain points should be noted about our two measures of net investment. In the first place, they are both free from the comparison of the known with the unknown which Professor Hicks sought to avoid by adopting the Swedish definition. But, while the *ex-ante* measure still remains a measure at the same time of the expected change in the value of the total capital stock over the period, the *ex-post* measure is no longer adequate as a measure of the change in the value of the total capital stock which actually takes place. Our subjective analysis has enabled us to by-pass the 'obsolescence' caused by mere additions to capital with constant techniques, but it has not enabled us to avoid the obsolescence caused by changes in technique. In fact, it is more than likely that an improvement in technique may, by rendering some of the existing capital obsolete, actually reduce the total value of the capital stock. However, we should not, in the present context, take this obsolescence too seriously, and we should particularly beware of concluding that as a result of it the position of the community has changed for the worse. Any fall in the value of capital will be more than outweighed by the enhanced value of what Professor Hicks has called 'human capital', by the increased opportunities for enjoyment which are now open to the individuals of the community.[1]

[1] Cf. Hicks, *Value and Capital*, p. 178.

This is as far as it is profitable to continue the present line of analysis. Indeed, the realist may think that we have gone too far already, though I think myself that we have learnt something during the course of our exploration of the subjective road, especially perhaps concerning the nature of obsolescence. There are certainly serious objections to a subjective approach, as Professor Hicks has already suggested.[1] Thus, our measures break down as soon as we try to allow for uncertain expectations. Nor, being subjective measures, would it ever be possible to use them in practice. Moreover, even if they could be calculated, it is not easy to see to what use they could be put, unless our dictator might wish to include them as part of his annual issue of propaganda.[2] Indeed, it might well be complained that we have so far been struggling to obtain measures of net investment and of net income without having any clear idea of why we want them. It is in the hope of remedying this defect that I now turn for guidance to Professor Hicks's 1940 article.

3. In very broad terms, what we require are two measures: a measure of net social income for use as an index of welfare possibilities, and a measure of net social output for use as an index of production possibilities.[3] Let us consider each of these in turn. With regard to the former, it will be remembered that at the outset Professor Hicks had doubts as to whether net investment, or its counterpart in saving, should be included in net social income at all.[4] And the point has given trouble ever since. For example, Dr. Little argues—rightly, I think—against the view that a decision to save should for this purpose be regarded as a decision to purchase investment goods;[5] for the two kinds of decision are in real life taken quite independently. But there are serious objections also, some of them provided by Dr. Little himself, against his own view that savings as such should be included in the net social income. In the first place, their inclusion renders inapplicable the index-number theory on which the use of the measure as an index of welfare possibilities is based. Moreover, once we concede, as we must if we are to make any sense of the inclusion of

[1] 1942 article, pp. 174–5.

[2] It might be thought that the *ex-ante* measure would help the dictator in deciding how much investment he should plan. But in fact he must already have solved this problem before he can use the measure. In any case, such decisions are—in theory at least—more easily taken at the margin than in the aggregate.

[3] I have in each case introduced the word 'possibilities' in deference to later discussion of the 1940 article. I shall not of course be concerned here with the wider questions of the validity and significance of index-number theory, but this may be the appropriate place to acknowledge the debt I owe to Professor Hicks for reading a paper recently in Oxford in which the 1940 arguments were reconsidered. Although I have not knowingly anticipated any of his results, nevertheless his paper gave a sense of direction to my earlier scattered thoughts on the subject of the present article.

[4] 1940 article, p. 123.

[5] *A Critique of Welfare Economics*, O.U.P., 1950, pp. 227–8.

savings, that the individual obtains current satisfaction from any addition he makes to his wealth, it is difficult to see why we should stop there. Ought we not to hold also that he obtains current satisfaction from the mere possession of wealth? The very distinction between income and capital, let alone the distinction between net saving and a capital gain, seems incapable of being drawn; and we should waste our time in trying to draw it. Instead, we should decide quite firmly to exclude net saving from our welfare measure, whether or not we then take Dr. Little's advice and replace the term 'welfare' by 'standard of living'.

The production approach seems to offer more hope; and it looks very much as if we shall be able to obtain a sensible measure of net investment and thence of net social product. For, so long as we are concerned only with production possibilities, it seems that we do not have to consider at all the question of obsolescence, which is so troublesome in other connexions. We can confine our attention entirely to physical wear and tear. This means that Professor Pigou's measure of net investment, although it was put forward for another purpose, is likely to provide the answer: for his is a measure of physical capital, obtained by weighting each type of capital equipment by its value at the end of the period.[1]

Indeed, I think we may regard Professor Pigou's measure as a good first approximation to what we require. But there are nevertheless certain theoretical difficulties about it. In the first place, we should, I think, want to exclude from our measure of wear and tear any physical depreciation which takes place not actually as a result of the production process. This Professor Pigou's formula does not do. Secondly, we may ask whether the weights he uses are the relevant ones to our present purpose. Ought we not to weight each type of capital equipment by its cost of production rather than by its value?[2] Thirdly, why should we be involved at all in a special index-number problem in the case of capital goods, when we are not in the case of consumption goods?

The distinction between investment and consumption goods which is implied by the special treatment given to the former is, I believe, partly a superficial one. The ordinary theory of index numbers as applied to con-

[1] Pigou, op. cit., pp. 273–4.

[2] In one not unimportant respect, Professor Pigou's use of end-year 'value' weights tends to correct the bias imparted by the failure of his measure to distinguish between the wear and tear which is the result of the production process and that which is not. As I see it, we should want to regard the physical depreciation represented by the actual scrapping of *obsolete* equipment as part of the latter; for, although no doubt the scrap *material* is again used in the production process, the scrapping of the *machine* is strictly a result of obsolescence, not of the production process. Now, the use of end-year 'value' weights conveniently brings it about that this kind of physical depreciation is given very little weight. For, since the type of machine in question is quite obsolete at the end of the year, its value will only be its scrap value. And this is just about the weight it ought to have, if it is to be regarded as a raw material and not as a machine when it is finally used up in the production process.

sumption goods tacitly assumes that prices remain constant *within* each period, while in any period in real life this is not of course the case. The assumption of constant prices therefore partly vitiates the application of the theory, though hardly to a serious extent; and the assumption is clearly a convenient one. On the other hand, no such assumption could in general be made in the case of capital goods, since the change in the price of capital goods was more often than not regarded as part of the problem under consideration. But it is not part of the problem so long as we are concerned only with production possibilities. As Professor Hicks already pointed out in 1940, consumption and investment goods are all on a par when looked at from this point of view.[1] Our particular problem can be tackled quite straightforwardly. There is no difficulty at all in obtaining a measure of the gross social output; and, in order to arrive at net social output, all we have to do is to regard physical wear and tear, as it occurs in the production process, as negative production, measuring it by the cost of replacement. This then is the solution to our, albeit now limited, problem—a solution already reached by Professor Hicks in 1940.

But is it really a solution? There is still one small point left to trouble the conscience. At the end of the preceding paragraph, we spoke, perhaps too readily, of the 'cost of replacement'. What exactly do we mean by the phrase? We cannot mean the cost of restoring each piece of equipment literally to its former condition. For this might be a very expensive, even an impossible thing to do. In real life wear and tear is more usually made good by the production of a new piece of equipment, even if still one of the same type; and the production of a new piece of equipment may do rather more than merely make up for the wear and tear in any period on an old piece. We have, then, to decide how much of the cost of production of a new piece of equipment should be regarded as making up for the wear and tear on another piece, and how much should be regarded as net investment. But this is the very problem we have been trying to solve, and it remains a problem even when we are concerned only with physical wear and tear. We may have succeeded in reducing the area of the circle, but we are still left on the circumference, with no prospect of ever being able to reach our goal in the centre.

The same difficulty arises when we try, as we suggested above, to replace the value weights in Professor Pigou's measure by cost of production weights. For, what is the cost of production of a partly worn-out piece of equipment? We can give no answer to this question unless we have first arrived at a measure of the cost of making good wear and tear. Lest we allow the impossible best to be the enemy of the good, we have in the end

[1] 1940 article, p. 123.

to fall back on Professor Pigou's measure, even though his weights are not those we should ideally want to use.

4. In conclusion, we may summarize the position in this way. The whole problem arises because of the heterogeneity of capital, even of capital used in the production of a single commodity. This heterogeneity exists for two reasons: first because the introduction of new techniques alters the type of equipment; secondly because physical wear and tear alters the condition of equipment. The foregoing discussion has certainly not led us to any solution of the problem. Rather, we have shown that it is insoluble—in itself perhaps not very startling news. Nevertheless, what we have done, I think, is to make clearer what are the reasons why no solution is possible; and here we may have succeeded in casting doubt on some preconceived opinions. I myself certainly believed, even up to a late stage in the preparation of this paper, that the problem of physical depreciation was relatively manageable and that the obsolescence caused by the introduction of new techniques was the stumbling-block. It turns out that the opposite is the case.

We suppose at the outset that we require a valuation of aggregate net investment for the construction of three indices: an index of current welfare possibilities, an index of future welfare possibilities, and an index of production possibilities. We rule out the first use in deciding that neither investment nor saving should properly be regarded as contributing to current welfare. The second use seems to have more justification; and it would be idle to deny that innovations are at the root of most of the difficulties of valuation in this connexion. At the same time, however, they render nugatory the whole purpose of the valuation. For, what we require for the construction of the second index is not so much a measure of net investment as a measure of total capital. When innovations are taking place, net investment no longer represents the change in the value of total capital. Nor indeed would the latter be of much significance, since it does not take into account the very important changes in the value of human capital.[1] We are left then with the third use as the only one for which a valuation of aggregate net investment can be justified.[2] Here, it is not obsolescence but physical depreciation that presents the difficulties—difficulties which cannot entirely be overcome. Our inability to overcome them can be laid at the door, I think, of a single feature of economic life: people do not, in general, *produce* partly worn-out machines; they produce new machines and then wear them out.

THE QUEEN'S COLLEGE,
OXFORD.

[1] Moreover, it would be at the mercy of mere changes in expectations, and would involve the comparison of the known with the unknown.

[2] With some emphasis on the word 'aggregate'.

THE ROLE OF INVESTMENT IN TWO FRENCH INFLATIONS

By HAROLD LUBELL

SINCE the war, France has gone through two serious price inflations. The first ran wild from the liberation through the end of 1948. The second started in mid-1950, at the outbreak of the Korean war, and came to an end in mid-1952, connected in the public eye with the 'Pinay experiment'. Both inflations were closely linked with the volume of investment expenditures in France, the first under the impact of post-war reconstruction and development needs, the second as a reaction to the strained international situation and the prospects of a domestic armaments boom. The present paper argues that the first inflation was due more to the exceptional circumstances of the times than to faulty government economic policy, that the second can more legitimately be blamed on the lack of action by the government authorities, but that neither of the two past inflations is an excuse for the lack of a policy of economic expansion in a period of stagnation.

The pre-1949 French inflation differs from that of 1951 both because of a basic difference in the type of resource problem facing the country and because of the different type of investment pressures on the economy. In the early post-war years, the principal problem was one of restoring production, with the government investment programme facing a deliberate choice between equipping the economy with increased productive capacity for the future and concentrating on an increase in production for immediate consumption. In 1951 it was apparently a choice between rearmament and domestic investment, but in retrospect it became clear that inflation had occurred without a significant amount of rearmament, as the result of a boom in private investment that went on completely outside the scope of the government's investment programme. However, despite the evident differences between these two situations and between both of them and that of 1954, the French Government seems to have become convinced that an inflationary threat is part of the structure of the French economy, a conviction that has prevented it from taking any positive economic policy step for several years.

The Post-war Inflation

France faced the problem of all the culturally advanced but economically backward countries since the middle of the nineteenth century, that of greatly increasing the volume of consumer goods when a prerequisite

was the prior creation of a highly organized and effectively functioning industrial plant. There was, however, one essential difference between the case of France after World War II and the cases of Germany and Japan in the latter part of the nineteenth century and Russia after World War I. In Japan and Germany the process of rapid economic development had taken place during the relatively peaceful conditions of the post-1870–pre-1914 era, where industrious populations started with an acceptable standard of living; in Russia, the standard of living, if not acceptable, was at least habitual. In France the reconstruction effort took place when the standard of living was far below what was considered as normal. A reconstruction and investment programme therefore found itself faced with the strong possibility that current investment not only prevented an immediate increase in current consumption, but also prevented a return to accepted pre-war levels of consumption.

The Monnet Plan, organized to direct the French investment programme during the first year after the liberation, and put into operation at the beginning of 1947, attempted a co-ordinated effort to reconstruct and modernize the damaged capital equipment of the country. A good deal of reconstruction was already taking place: a major effort was being made to get the damaged railway system back into operation by repairing destroyed bridges, road beds, and traffic installations and by finding and replacing damaged and lost rolling stock; coal-mines were put back into operation partly through the use of German prisoners of war; and so on. The level of investment undertaken in the immediate post-war period was high. In 1938 only 3 per cent. of the gross national product was devoted to investment, and over one-half of this went for gross capital formation by government. In 1946 and 1947 gross domestic investment amounted to 15 per cent. of the gross national product on the U.S. definition;[1] in 1948 it was reduced to 12 per cent.[2]

As might be expected, a good part of the investment in 1946 was in the form of inventories as enterprises, but the rate of inventory accumulation declined greatly in 1947 and 1948, partly as a result of the curtailment of bank credit for such use at the insistence of the government. Gross capital formation by government itself was quite small throughout the post-war years. On the French definitions of gross investment and gross national product, the investment effort took between 18 and 19 per cent. of the gross national product during the years 1946–9.[1] On either definition, the French economy was extending an effort five or six times as great as during

[1] On the U.S. definition, current repair and upkeep are excluded from the gross concepts. See Richard Ruggles, 'The French Investment Program and its Relation to Resource Allocation' (E. M. Earle, ed., *Modern France*, Princeton, 1951), p. 380.

[2] Economic Co-operation Administration, *National Accounts Data Book 1951* (Washington, 1951), Tables 1 and 3 for France.

the final year of the inter-war period, out of a gross national product smaller in real terms than in 1938. The most important question for French economic policy after the war was whether or not the economy could stand the strain of such a high level of investment at a time when it was recovering from the shocks of war and suffering from a runaway inflation.

The policy of the Monnet Plan was to put through its investment programme at any cost in order to create, as soon as possible, sufficient productive capacity to guarantee a higher standard of living in the future. Starting from its premiss of the indispensability of its programme, it claimed that the inflationary situation permitted only two types of policy. The first was one of specific controls: price control, materials allocation, rationing, &c. The second was a policy of global monetary controls: increased taxes and credit restriction. However, the specific controls had an unsuccessful history in France after the liberation: serious efforts were made to control prices during only three months of the period of shortage, from January to March 1947, when the Blum Government imposed two successive price rollbacks; materials allocation was taken out of the hands of the war-time *comités d'organisation*, tainted with the corporate-statism of the Vichy régime, only to be given over to their alter-egos, the *syndicats patronaux*, which were the new employers' organizations in each sector; the compulsory collection of meat and the rationing of bread had been abandoned during 1945. The government therefore turned to global controls. The imposition of credit controls at least stopped some financing of speculative inventory accumulation by bank credit, but was evidently not able to stop the inflation. An exceptional levy of 150 billion francs, including a forced loan and the recall of all 5,000 franc banknotes, balanced the budget in the first half of 1948. Otherwise, budget deficits continued so that the fiscal policies carried out had no deflationary effect on incomes.

In support of the Monnet Plan's choice between investment and consumption, it can be argued that even though consumption was at a dangerously low level in the early post-war years, the investment programme did not hinder an increase in consumption, and, on the contrary, was in some measure indispensable for the realization of an increase in consumption even in the short run. This is borne out by the available figures on real aggregate consumption by commodity (which, however, include purchases by government as well as purchases by consumers), in 1938 and in the post-war years, measured in 1938 francs.[1] In the first half of 1946 total real consumption at an annual rate was only 80 per cent. of

[1] Economic Co-operation Administration Special Mission to France, *France Data Book* (Paris, April 1950), Table N-3.

1938; in the second half of 1946 it rose to 88 per cent., but it fell off again in the first half of 1947. The consumption items hardest hit by the period of shortage were food and durable goods, both of which were down to 75 per cent. of 1938 during the first half of 1946. Although consumption in France was much lower relative to pre-war than in the other Western European countries except Italy and Germany,[1] the first half of 1947 saw a substantial approach to pre-war standards in all categories except food, semi-durables, and housing.

The most serious shortage was therefore of food, and there were also significant shortages of leather and perhaps of clothing. In these cases, investment did not compete for domestic resources in the short run, nor did the investment programme directly reduce their output. Agriculture depends on the weather; the production of textiles on imports of raw materials. There was, of course, direct competition for resources by way of the balance of payments; with regard to the shifting of domestic resources, this was not the case. Admittedly a cut in the investment programme might have eased the food situation if an immediate increase in industrial consumer goods had resulted and induced the farmers to hoard less food. However, the real difficulties were a series of droughts and bad harvests rather than the traditional habits of the peasantry.

Another severely felt shortage was coal for household use, and to some extent this was competitive with the investment programme since the coal shortage was felt everywhere. In any but the shortest view, however, the coal shortage demanded a rapid increase in investments in the coal industry rather than a curtailment of investment. Housing was obviously in a different situation, since it was an area distinctly competitive for resources with investments in industry and elsewhere. Both construction materials and building workers were in short supply throughout France, and an active decision was taken to limit housing construction until the general supply situation had been eased by expansion of the output of building materials, steel, &c.

The production of consumer durable goods developed at a surprisingly rapid rate, considering how directly they compete with investment goods for resources. Despite the reconstruction and investment effort, consumer durables were within 3 per cent. of the 1938 level by the second half of 1947 and exceeded the pre-war level by 10 per cent. in the first half of 1948. Thus, the type of consumer goods apparently most likely to be displaced by investment rose rapidly, concurrently with the investment programme, while consumer goods not directly competitive in terms of domestic resources rose more slowly. It would therefore seem that a smaller amount

[1] United Nations Economic Commission for Europe, *Economic Survey of Europe in 1949* (Geneva, 1950), p. 26.

of investment would not necessarily have improved consumption levels any more rapidly than they were improved.

The extent of the inflation, combined with the failure of the recommended measures of control, nevertheless led to demands that the investment programme be cut. In France they were voiced by orthodox financial advisers of the Ministry of Finance and some of the more conservative members of the Conseil de la République (the French Senate). One of the best criticisms by an outside observer was written by R. V. Rosa,[1] who argued that the continuance of the inflation after 1947 was due to the French attempt to impose an ambitious programme of new investment on an economy unwilling to accept the sacrifices in voluntary or compulsory saving that such a programme required. The only way of stopping the damage done by the inflation was therefore to curtail the investment programme and to aim at post-E.R.P. viability at a lower level of total output and a lower standard of living than hitherto desired.

In December 1948 the statisticians of the Monnet Plan estimated an 'inflationary gap' of 475 billion francs for the year 1949. By June 1949 the wholesale price index in France reached its low point after a steady decline from January 1949, and the first post-war inflation was over. The controls were inadequate, the taxation system was ineffective, the investment programme had escaped its parliamentary battles almost uncut, but the unexpected had occurred and the inflationary crisis was over.

The essence of the explanation of the end of the pre-1949 inflation is that the 'process of production' requires time. There are two elements of waiting to be considered in an economy starting production more or less from scratch.[2] The first is the time necessary to build the new capacity which is to be brought into use in the production of finished goods. The second is the time necessary to get products out of the process after putting in the materials. The explanation of the unexpected end to the galloping French inflation would seem to lie in this second element of economic waiting. Although it appeared that the attempt to advance beyond the old normal levels of production was preventing even the attainment of these former levels in the short run, while the inflation ruined the economy and its chances for the long run, the shortages at the consumer stage came to an end when the raw materials in the production process started coming out as finished goods. The 'ripening of the production process' thus saw the arrival on the market of products which had been passing through the various stages of production since the start of the post-war recovery. From the increase in stocks and work in progress at all the stages hidden from

[1] R. V. Rosa, 'The Problem of French Recovery', *Economic Journal* (June 1949).

[2] P. Uri, 'L'évolution de l'économie française jugée par la théorie moderne', *Réalités* (Paris, December 1950).

the general view, there eventually came an abundance of final goods sufficient to fill the shelves and stop the constant price rises.

The end of the first post-war inflation brought with it a small recession which in its turn caused great concern to the economic policy makers in France. Within 18 months, however, a new French inflation began.

The Post-Korea Inflation

The Korean conflict set off a price inflation in most European countries and in the United States, but it showed its greatest severity in France. In France the upward trend in prices began in July 1950 and continued until January or February of 1952. The initial impulse of the price rise came from abroad, through the sharp increase in raw material prices which first affected French domestic prices at the wholesale level and then spread to the retail level. By April 1951 the imported raw-materials sector of the French wholesale-price index was 94 per cent. above the June 1950 level, while the total wholesale-price index was 36 per cent. above. In January 1952, when the general wholesale-price index reached its peak 47 per cent. above June 1950, the imported raw-materials price index was down to 40 per cent. above the June 1950 level. Retail prices (as measured by the index of consumer prices in Paris) reached a peak in February 1952 at 39 per cent. above the June 1950 level. The forces which made the French inflation the worst in Europe and caused it to continue well after prices of raw material imports started to fall were, evidently, peculiar to France.

The peculiarity of the French inflation in 1951 lies in the coexistence of an investment goods boom and a consumers goods recession during the period of a general price rise affecting both investment goods and consumer goods. The general recession of the early part of 1950 had made itself most prominently felt as a decline in metals production and construction activity, while production in all other industries continued to rise. The upswing after Korea exerted its major effects on these previously slack industries, although consumer goods production also rose during the latter half of 1950 and through the first five months of 1951. After that point, however, consumer resistance to inflationary prices resulted in declines in the production of consumer goods, without preventing further price rises. Thus, by May 1951, the production index (on the base June 1950 = 100) for equipment goods excluding construction was 116 and for consumption goods 117. By January 1952 the index for equipment goods was 124 while the index for consumption goods was down to 113.[1]

Part of the sharp increase in prices which occurred in the last part of

[1] Mutual Security Agency Special Mission to France for Economic Co-operation, *France Data Book* (Paris, March 1953), Tables D–1, D–2, and H–1.

1951, after a small price decline during the summer, was due to governmental action permitting price increases for certain basic commodities (agricultural products, steel, coal, electricity and gas)[1] and a 15 per cent. increase in the legal minimum wage. The essential element in the inflation was, however, the investment behaviour of private producers.

Although a large-scale armaments programme never came into being during the inflationary period, the expectation of rearmament led to the great increase in private investment activity reflected in the production indexes. Investment expenditures financed out of public funds declined by 5 per cent. in current franc terms from 1950 to 1951, and by perhaps 25 per cent. in real terms, taking account of the intervening price rise. On the other hand, privately financed investment more than doubled in current franc terms during the same period. Private construction activity rose by some 60 per cent., in real terms, over the peak previously reached in 1949, while the installation of machinery and equipment by private firms rose by some 40 per cent. The increase in private housing construction, far lower than desirable on an absolute basis, was brought about partly in response to a government programme of construction bonuses. The larger part of the private investment boom was, however, in conflict with government policy, since the pressure on resources of private investment demand resulted in the effective reduction of the official investment programme far beyond those expected from the cuts in government reconstruction and investment expenditures imposed by Parliament.

The theoretically troublesome aspect of the boom and inflation was the lack of demand for consumer semi-durable goods. Demand for food remained high and exerted a constant upward pressure on agricultural prices. Unlike the earlier post-war inflation, however, consumer demand was not directed towards textile and leather goods, whose production fell during the last seven months of open inflation in 1951. The obvious sort of consumer good which might be expected to absorb excess purchasing power was simply not wanted at going prices.

A desirable result of such a state of affairs would have been a reduction in prices, in an attempt to promote sales. Instead, production was reduced. Producers and merchants were evidently counting on further inflation to permit them to clear their shelves without a current franc loss. A good part of the onus of the continuation of this situation lies with government credit policy which permitted the continued extension of bank credit to the textile and leather producers because they were in difficulty. Instead of

[1] Despite the general breakdown of price controls, certain agricultural prices, such as wheat, are still decided by agreement between the government and the producers. The energy sectors, except for petroleum refining, are nationalized and therefore directly responsible to the government, while steel prices were determined on an avowedly cartel basis before the institution of the Coal and Steel Community.

being forced to liquidate stocks, the merchants and producers were able to carry overpriced inventories until the inflation had come to an end.

In retrospect, it is apparent that the inflation stopped in February 1952. Politically, it nevertheless created a parliamentary crisis through May 1952, when the Pinay Government took office with a platform of 'no increases in taxes' and a general 5 per cent. price reduction campaign agreed upon by the Premier and the big department stores. Until the advent of the Mendès-France Government, the fear of a renewed inflation obstructed the clearcut creation of French government economic policy.

There are several reasons why the second inflation came to an end, other than the self-evident one that most inflations eventually wear themselves out. The most important reason was that there was no objective cause for the inflation other than the conviction of most Frenchmen that prices would continue to rise: the rise in world prices, which had set off the inflation in France had been reversed some 10 months previously; the plans for a military programme, in whose anticipation the private investment boom came about, were cut down during the inflationary situation before coming to fruition; the recession in the consumer goods industries resulted in some unemployment and declines in income. At the same time the propaganda aspect of the Pinay programme was effective to a certain extent. The 5 per cent. price-lowering campaign had its part in convincing people that prices might not rise, particularly since it coincided with the normal spring price declines in agricultural products. A good deal of publicity was given to cuts in the investment budget whose purpose was to curtail investments by the nationalized enterprises, and scheduled price increases by some of the nationalized enterprises were cancelled. As a result, the Pinay 'experiment' succeeded in forestalling demands for new wage increases, which, if obtained, might at that point have stimulated new price rises.

The attempts by Pinay to curtail the government's investment programme in current franc terms turned out to be in vain. Although the investment budget had ostensibly been cut to 923 billion francs, final expenditure figures turned out to be 1,000 billion francs: the nationalized enterprises financed their scheduled programmes by short-term bank borrowing and maintained the planned rate of investment. It is therefore clear that there was no reduction in public investment to reduce pressure on resources in 1952. Private fixed capital formation, on the other hand, started to decline in the last two months of the year. Over the year, however, the real volume of fixed capital formation in enterprises declined from 1951 to 1952 by some 7 or 8 per cent., since the 5 per cent. rise in investment in current franc terms was much more than offset by the 14 per cent. price rise from 1951 to 1952.

One of the key factors in ending the inflation was a divergent movement of consumption and production of consumer goods. Real consumption expenditures rose from 1951 to 1952, although the production of consumer goods other than food declined. What seems to have occurred was a sharp curtailment of orders to producers while wholesalers and retailers concentrated on getting rid of their overpriced inventories without lowering prices. An important alleviation of pressure on resources therefore resulted from a sharp decline in the rate of inventory accumulation by the end of 1952.[1]

The Official View of Inflation

The 1951 inflation led to the development of an official theory of French inflation which may be fated to interfere with French economic policy for a long time to come. The Ministry of Finance's 'Report on the National Accounts for 1951'[2] contains a chapter, by Simon Nora (one of the Mendès-France brains trust), developing the notion that France's past inflationary tendencies are due to 'structural rigidities' peculiar to France. Nora generalizes from the 1951 combination of inflationary pressure on the equipment goods sector and recession in the consumer goods sector, to a general statement that the French economy is incapable of adjustment to shifts in demand which put pressure on resources in any sector. The conclusion drawn by Nora is that a concerted investment programme, cushioned by anti-inflationary controls, is necessary to break down the structural rigidities in order to permit expansion of the economy. It is an argument that had its most prominent use in combatting further expansion of the French military effort in 1952 and 1953. As might have been expected, it unfortunately also led to pressure to reduce the investment programme, since French Finance Ministers have a tendency to concentrate on short-run inflationary prospects—even when the *conjoncture* is deflationary—rather than on long-term expansionary possibilities.

The trouble with the rigidity argument is that it may lead to a dodging of the real policy issue, which is the responsibility of the government's financial and fiscal actions for the inflationary or deflationary state of the economy. In 1951 the government failed to curtail excess private investment demand and permitted the extension of credit to bail out producers who were unwilling to lower prices in the face of consumer hostility. Putting the blame on the 'structure' of the economy permits the Finance Minister to avoid trying to carry out an expansionary policy during a

[1] Mutual Security Agency Special Mission to France for Economic Co-operation, *France Data Book* (Paris, March 1953), Table A–1.

[2] Service des Études Économiques et Financières du Ministère des Finances, 'Rapport sur les comptes provisoires de la nation de l'année 1951', *Statistiques et études financières: Supplément finances françaises*, No. 18 (Paris, 1953).

period of stagnation on the ground that an effort to stimulate the economy will be inflationary. During 1953 private investment activity fell off from the previous year's levels in spite of a recovery in consumer goods production, while expenditures out of the government's investment budget were at the same level as in 1952. Price stability was maintained, but it would probably have been equally well maintained if combined with a more energetic government investment effort. One essential trouble with the French economy is that action is taken to stop an inflation once the inflation is at an end, and the past inflations are then used as an excuse for not changing to a policy of stimulation when the latter is required.

FALK PROJECT FOR ECONOMIC RESEARCH IN ISRAEL,
JERUSALEM.

INDUSTRIAL RELATIONS IN NINETEENTH-CENTURY BIRMINGHAM

By ALAN FOX

By the middle of the nineteenth century it was already becoming a commonplace of economic and social observers that the pattern of industrial relationships in Birmingham had certain peculiar features which marked it off from other industrial centres such as Manchester (with which it was frequently compared). Any reference to these peculiarities was invariably linked with a reference to the prevailing small scale of Birmingham industry and the proliferation of little workshops with their close personal contacts between the small masters and their workmen. The *Edinburgh Review* remarked in 1859 that 'Birmingham is little troubled with serious disputes between employers and employed, chiefly because the manufactures of the place are of a kind requiring incessant ingenuity and novelty, and therefore occupying individual minds in a way defying organization [of labour]. Masters and men consult about a new device in jewellery, or for domestic or decorative purposes, and each piece of work is paid for on its own merits or under its own conditions. Where hundreds of masters are employing thousands of workmen who are daily aiming at the production of something new, no uniform rate of wages can be established.'

J. T. Bunce, a Birmingham notability, looking back in wistful retrospect from the more impersonal days of the early 1900's, wrote that 'In the ordinary workshops . . . masters and men often worked at the bench together, and so were comrades rather than masters and servants. Ranks were not so sharply divided as to exclude real companionship and intercourse. . . . It was not uncommon for workmen to address their employer . . . by his christian name.'

Others were not slow to relate this industrial pattern to Birmingham's political atmosphere. Cobden, commenting in 1857 on Bright's first election for Birmingham, was emphatic that 'the social and political state of that town is far more healthy than that of Manchester; and it arises from the fact that the industry of the hardware district is carried on by small manufacturers, employing a few men and boys each . . . whilst the great capitalists in Manchester form an aristocracy, individual members of which wield an influence over sometimes two thousand persons. . . . There is a freer intercourse between all classes [in Birmingham] than in the Lancashire town, where a great and impassable gulf separates the workman from his employer.'

Foreign visitors, too, invariably commented on the peaceful tenor of relationships and went away sighing their envy at the admirably flexible

nature of the local labour market. And most of all, employers in other industrial centres, less felicitously endowed in these matters, were apt momentarily to break off their struggles with powerful and uncooperative trade unions to stare across at this Eldorado and bemoan their luck. For Birmingham was, above all, a centre where trade unionism was halting in its growth and weak in its striking-power. Sidney and Beatrice Webb, in their *History of Trade Unionism*, found Birmingham one of the black spots as far as trade-union weakness was concerned. Relating the ascertained number of trade unionists to local population in all the relevant counties in England and South Wales, they put Birmingham fifth from the bottom in a list of eighteen, with a percentage of 4·1 compared with Northumberland at the top of the list with 11·2 per cent. The Webbs, rarely tentative in their assessment, had no hesitation in attributing this to the 'small master' pattern of industrial organization that prevailed there.

This is indeed a large part of the answer, but unfortunately it has lent itself to oversimplified interpretations. Thus the prevalence of industrial peace has often been attributed merely to the direct personal relationships between master and man, as if personal contacts are invariably amiable; whilst the weakness of trade unionism has frequently been regarded as solely due to the fact that only in big establishments with large groupings of workers can unionism thrive. These assumptions have hitherto appeared to be sufficiently convincing in themselves to discourage any closer examination, and even the Webbs stopped a good way short of a full analysis. But much remains to be said, not only of the industrial organization that shaped Birmingham's unusual industrial relationships, but also of the behaviour, tactics, and philosophy of the trade unions which arose in response to them, of the leaders they threw up, and of the process of change which overtook all three.

From all these points of view there is much that was common to the 'staple' trades of nineteenth-century Birmingham. Jewellery, gun-making, button manufacture, and the brass and other metal trades shared many significant characteristics, and so between them produced that 'temper' of relationships which was peculiarly Birmingham in quality.

The basic characteristics were the continued use of manual skills, an exceptionally wide variety of product and design, and concomitant with these an absence of opportunities for standardization, for economies of scale, and for mechanized production.[1] The industrial structure associated with these characteristics consisted in each case of a handful of bigger establishments with a labour force of a hundred or two, a larger number of workshops each employing perhaps up to fifty, and a vast uncharted fringe

[1] The material relating to industrial structure and organization owes much to Professor G. C. Allen's *Industrial Development of Birmingham and the Black Country* (London, 1929).

of very small workshops, domestic outworkers and 'garret-masters', employing up to five or ten workers, or perhaps only members of the family. Markets were highly competitive, with the competitive effort becoming more desperate as one moved lower in the structural hierarchy.

The differences between the factory owner on the one hand, and the workshop proprietor and garret-master on the other, lay not only in size of establishment but also in methods of industrial organization. Both the workshop owner and the garret-master were heavily dependent on a merchant intermediary (usually known in the domestic market as a 'factor', the term 'merchant' being reserved for the foreign trade) in a way in which the factory owner was not. Whereas the factory owner could provide his own working capital and deal directly with the banks when he needed credit, the shop owner and garret-master were dependent on the factor not only for distribution of the products but also for financial assistance during production and often for raw materials also.

This dominating position of the factor had also both encouraged, and been encouraged by, his emergence as a vital figure in the co-ordination of production. The absence of opportunities for standardized production and for economies of scale not only permitted but positively encouraged the small unit to thrive, since given the necessary co-ordination the small units could specialize in single processes or even sub-processes. It was the factor who co-ordinated the activities of these small specialized units to produce the finished article.

Sometimes, in the interests of convenience and more effective co-ordination, or for the joint use of a common service of power, the factor gathered together some or all of his dependent workshop owners and garret-masters under one roof, supplying them with shop-room, materials, lighting, and power, for which they paid him a weekly sum. In such cases he became the 'employer' and they subcontractors, hiring and firing their own labour and being entirely responsible for it. Such were the origins of the few larger establishments as did exist.

Each 'subcontractor/workshop owner' was paid by the employer on straight 'piece-prices', and made his own arrangements with his employees (known as 'underhands') concerning their wages. The vital difference was that while he was paid on piece-prices, he paid his team on fixed day-wages. The more economically and efficiently he organized production the larger his profit. He usually sought this through elaborate subdivision of labour, each member of the team being kept for the most part on one or two simple skills which could be fairly easily acquired. Being under the maximum incentive to get his labour cheap, he often employed boys and youths, who could pick up a few of the easier skills and earn wages that were more than reasonable for their age. At the age of 16 or thereabouts

their rising cost in wages would prompt the subcontractor to discharge them and seek younger and cheaper labour. Those who found the opportunity to develop a knowledge of all the processes and skills which went to make up the skilled journeyman were comparatively few, and there was no regular apprenticeship system. There was consequently an unchecked influx of boy labour that could never be absorbed as 'skilled' workers, but which clung to the fringes of the trade, pulling down the wage-rates of older underhands, or in desperation joining the already crowded ranks of garret-masters to produce yet more shoddy work at starvation prices.

The relationships between the employer and his subcontractors were in essence the same as those between the factor and his dependent workshop owners and garret-masters. There was the same high degree of dependence of the one upon the other; the same wide diversity of product which reduced common interests between the small masters to a minimum and which made individual bargaining over piece-prices the rule; the same ease with which one small master or subcontractor could be played off against another.

Yet there were differences, and they were important ones. There was a certain psychological difference, for example, between the small master with his own workshop and at least an air of independence, however spurious, and the small master who as a subcontractor worked as one of many on the premises of an employer. In the former case, consciousness of his status as a master was likely to remain uppermost; in the latter this might be gradually giving way to a sense of subordination, in company with others, to the factory owner. This could serve as a basis for combination, but in some factories it was weakened by the close personal relationship which often existed between the employer and his subcontractors. Some subcontractors might have been with him all their working lives as he worked up his business. In such cases the personal attachment between the employer and a subcontractor of long standing might well be stronger than that between the subcontractor and his fellows, especially those who had not been with the firm for very long.

Thus organizational factors, the individual nature of the product, and personal relationships together made it difficult for the subcontractors to combine for the purpose of extracting better terms from the employers and factors.

If it was difficult for the small masters to combine against the employers, it was next to impossible for the underhands to combine against the small masters. They had little skill to protect them; and the same diversity of product and design that required individual bargaining between the subcontractor and employer imposed the same necessity upon the underhand *vis-à-vis* his subcontractor.

The underhand was perhaps doubly unfortunate, moreover, in having

no clear focus of grievance against which to direct his forces, such as they were. Any attempt to improve his position could be met by his subcontractor with the explanation that the prices he received from the employer did not permit a wage increase, and that the employer was in any case pressing him to take on younger and cheaper labour. Unless the team were a large one (some were up to forty or fifty), the underhand might well have some fellow-feeling with his subcontractor, since the profit made by the latter would certainly not be so sizeable as to alienate his men's sympathy. It seems highly probable that there were many cases in which the subcontractor was thought of by his men as more sinned against than sinning. In any event, the underhand's sense of grievance was certain to be divided as between subcontractor and employer, and incapable of being concentrated solely against either.

The organizational division into two levels of labour thus had profound effects upon industrial relationships. In the first place it split the ranks opposed to the employers and factors by setting up a conflict of interests between the small master and his underhands. And, in the second place, it greatly mitigated the bitterness of struggles waged by the master craftsman alone against the factor, by giving him the status of employer. The small master was never in any doubt that he had 'risen' from the ranks, and this sense of independent status (hollow in reality though it was) greatly influenced his relationships with employers and factors. While he might well feel bitter against a factor who took advantage of a vastly superior bargaining position to force down piece-prices, his deep conviction that in spite of all he was still 'his own master', employing his own labour and therefore not to be grouped with the undifferentiated mass of underhand day-wage labour, made it difficult for him to apply to himself any concept of 'class-exploitation'. He was unable to form consistent moral arguments based on any idea of exploitation of labour since he himself was an employer and cheap labour was one of his nearest interests. Even in depression, when struggles took place as to who should bear the major burden, the only real difference as he saw it between the factor and himself was likely to be that the factor was winning. Factors were men of a status admittedly higher yet similar in kind to his own. Disputes with them might well be brisk, they might even be protracted, but they could hardly become embittered with the missionary passion of a class-war. It was as difficult, in fact, for the small masters to form any consistent views of employers and factors as 'class enemies', as it was for the underhands to apply similar concepts to the small masters. The difficulties were heightened by the fact that nowhere in this pattern of status was there any impassable gulf. Just as the energetic and ambitious underhand could become a small master, so could the energetic and ambitious small master become a factor.

Since it was difficult to formulate economic grievances in 'class' terms, such grievances tended to focus on individuals, or on the notion of individual selfishness. There was a tendency to talk in terms of 'good' and 'bad' employers rather than to impute evil to classes, either social or economic. There is some evidence that this tendency was reinforced by the strong influence of local Nonconformity; to a preference for talking in terms of individual sin rather than of evils arising out of class structure. Certainly the Nonconformist influence, allied with political Liberalism, reached all sections of local society.

This emphasis on individual selfishness led to the emergence of scapegoats and handicapped any analysis which took account of structural and organizational factors. The 'middleman' or subcontractor or piece-master who 'sweated' his labour could all too easily be picked out as the origin of low wages and bad conditions. Contemporary opinion, having found its 'bad' employers, frequently looked no farther, leaving the factor to enjoy the principal benefits while evading the moral responsibility for their origins.

It is hardly surprising, given this industrial structure and organization and the relationships which grew out of them, that the growth of labour organization in the staple trades should be slow and its bargaining power weak.

Organization, when it did come, arose first among the subcontractors, and it was the boom of the early seventies which provided the stimulus. The boom, with its acute shortage of labour, had a galvanic effect on local labour as it did elsewhere. The Birmingham master craftsmen, usually so deficient in the community of interests which serves as the necessary basis for combination, found that at last they had something in common—scarcity value. And it was in this sudden access of strength that they found unity.

Small, shaky unions began to emerge among the subcontractors employed in the factories and larger workshops. They were immediately beset with difficulties. Were the underhands and the few day-wage workers employed directly by the factory owners to be encouraged to join or not? If so, union rules would have to ensure that underhands could not claim strike pay for striking against their piece-masters, who had formed the union in the first place and were subscribing most of the funds. What were union tactics to be when the abnormal labour scarcity had eased and the new-found common bond had melted away? Could the union hope to organize that jungle fringe of garret-masters struggling desperately for survival who broke down every protective fence that the better-placed craftsmen tried to erect?

It soon became clear that the circumstances of local industry were producing some strange alliances which could not be explained in the simple terms of 'capital versus labour' which might serve well enough for other industrial centres. The strongest allies of the subcontractors' unions were the larger-scale employers, who were able and willing to pay acceptable piece-prices and who looked to the union as a possible means of reducing the cut-throat competition. If the union could organize the garret-masters and small workshop owners, it could perhaps stiffen their resistance to those employers and factors who were exploiting their weakness and gaining thereby a competitive advantage.

The greatest enemies of the unions were these same garret-masters and the factors who dealt with them. The characteristic disposition of forces, therefore, was not that of 'employers versus workers', but of larger-scale employers and subcontractors on the one side, ranged against factors, smaller-scale employers, and garret-masters on the other.

Moreover it quickly emerged that even the subcontractors in the larger factories were a far from unified force. The fact that the larger-scale subcontractors were likely to feel closer in sympathies and status to their employer than to the small-scale subcontractor on the next bench with only two men and a boy, inevitably weakened union influence even where it professed its greatest strength.

The question of what tactics and weapons the unions were to employ quickly became preoccupations of the inexperienced leaders. The most astute of them, W. J. Davis, who built up the National Society of Amalgamated Brassworkers and was to remain its general secretary, with one short interval, until 1920, soon realized that craft unions in the Birmingham industrial context were in no position to extract concessions by coercing opponents with a display of strength. Not only was there the eternal spectre of the garret-master on the doorstep, offering the employer alternative sources of supply at cut prices; there was also the fact that where so many diversities existed it was difficult to find a community of interests sufficiently obvious and sufficiently wide to persuade an effective proportion of the men to take drastic action together.

If coercion through strike action was impracticable as a union weapon, so was the attempt to manipulate public opinion by evoking pity. Not that Birmingham was beyond the reach of humanitarian appeal—indeed there were abundant signs in Birmingham of Christianity being not only professed but practised. And in a society where the range of living standards was wide, but where social conscience was strong, considerable effects could be achieved by the more poverty-stricken groups through an appeal to pity. But any attempt to invoke pity with respect to the organized craftsmen of the local staple trades would have carried little weight and

might even in some cases have appeared slightly ludicrous, since such claims would be coming from artisans who, though poor, were not abysmally so.

Since coercion and the evocation of pity were impracticable weapons, the only alternative was the appeal to 'reason' and 'fairness'. Union leaders carefully nurtured the approach of the honest, sincere, straightforward Englishman appealing to the employer, as one reasonable, humanitarian gentleman to another, to admit that his piece-prices, or a particular piece-price, insulted the skill of the subcontractor concerned. Evidence would be advanced purporting to show that such prices barely covered his underhands' wages; that the underhands' wages were in any case gravely inadequate and ought to be increased, but that this could only be done if the employer raised his piece-prices; that they compared badly with prices paid for the same work by similar firms; were disproportionate to prices paid for other work by the same firm; barely allowed the workman to keep his family in Christian decency; earned the employers a bad name; compared so badly with prices paid in other brassworking towns and in other industries as to earn Birmingham an evil reputation, &c.

The accumulated weight of arguments such as these might be far from negligible if delivered by union leaders who were prominent in all forms of local (and sometimes national) public life. And in Birmingham the links between trade-union office and public office (on local councils, school boards, and boards of guardians) were very close, especially from the eighties onward.

But whatever the tactics available, local leaders were only too aware of the difficulties which the employers faced, however personally willing they might be to concede union demands. The cost structure of local industry, with its high proportion of labour to capital, meant that wage-rates tended to be very sensitive to changes in selling-price, and this drove union leaders to focus their resentments as much on 'the market' as on the employers. They were all the readier to do so since they themselves had direct personal experience of the exigencies and pressures of the 'employing' function. For these reasons they were only too ready to abandon the attack on the employer, and instead try to persuade him to join them in an attack upon the consumer.

In its simplest manifestation this took the form of seeking to convince the employer that an increase in piece-prices need cost him nothing if he raised his selling-prices by the amount necessary to cover the increase. It would be pointed out that the public, in its careless ignorance, paid whatever the lowest price happened to be without troubling to consider whether such a price paid the employer a 'reasonable' profit and the workman a 'fair' wage. The defenceless workman, the struggling employer, were at

the mercy of this irresponsible buying public. Yet, the argument continued, did the public really care whether it paid 1*s.* 2*d.* or 1*s.* 3*d.* for an article? It did not. The only reasonable course, therefore, was to fix the price at the level which gave the employer his 'reasonable' profit and the workman his 'fair' wage.

At the root of the argument, of course, lay a deep hostility towards competition, the prime enemy of so many trade-union leaders of Birmingham. The impracticability of controlling entry into a trade and restricting its output brought the competitive principle into their acute disfavour, and the buying public came to appear to them as the source of their main troubles. The natural outcome of this conviction was the idea that employers and workers should come together in an alliance for mutual protection against this common menace. It was an idea which came easier to men in the industrial context of Birmingham, where relations between employer and craftsman were close and personal, and where the craftsman himself could without much difficulty acquire 'employer' status, than in a context where relations were impersonal and soured with the bitterness and distrust of rigid social stratification.

The employers' reply to these overtures was usually to point out the impossibility of controlling the many small masters who between them were responsible for much of the destructive price-cutting. An increase of selling-prices by one or a few employers alone would benefit only those firms who refused to follow suit, since trade would immediately pass into their hands. If the union really wanted to help the trade it should force up wages in all the small workshops, thereby eliminating the unfair competition which prevented the employers paying their men the piece-prices they would like to pay.

But despite repeated rebuffs, union leaders continued to urge 'capital and labour to unite for the protection of British industries', and gradually evolved the philosophy of a capital-labour alliance which would suppress competition, maintain and, if possible, raise selling-prices, and share the resulting increments of profit.

It received its fullest practical expression in the Alliances which sprang up in Birmingham during the nineties among different branches of the metal trades. The initiative came, characteristically enough, from an employer who had previously been a craftsman, but there was strong support from union leaders, particularly Davis. In each case a Conciliation Board was set up, representing both employers and men, which, in addition to the normal functions of such boards, administered the special features of the Alliance.

The peculiar feature of the Alliances was the scheme by which the unions concerned constituted the main method of enforcing the Alliance's price

policy. In return for the 'closed shop' and a certain percentage bonus on wages which accompanied every unit increase in price, the unions guaranteed to call out their men from any firm which sold below the price fixed by the Alliance Board. Accounts of the practical application of this guarantee can be found in the local press of the time. The Alliances themselves were discussed freely and openly, and their sponsor, E. J. Smith, published a book which he clearly hoped and anticipated would spread the light. Both he and union leaders such as Davis sincerely believed that the key had been found to the ideal industrial system. None of the Alliances survived the turn of the century, though in the meantime they achieved some success; enough anyway to confirm Davis in his philosophy.

Davis was throughout the whole of this period on the Parliamentary Committee of the T.U.C. and in close touch with the rest of the trade-union movement. There may well have been a certain piquancy in his relationships with other union leaders whose attitudes had been shaped by a very different kind of industrial context.

Whatever the attitudes and tactics adopted towards the employers, however, they were incidents in a struggle on one front only, and it was the curse of a union of piece-masters that it had to maintain its defences on two. Just as the employers had to be withheld from squeezing down piece-prices, so the underhands had to be prevented from pushing up their day-wages. For most of the time this presented little difficulty, since the weakness of the underhands' position was notorious. But it did mean that union officials were directly interested in maintaining a free inflow of young boy labour in order to maintain this weakness. When, in 1875, the Royal Commission on Factory and Workshops Acts held a number of inquiries to test opinion on the proposal to raise the minimum age at which full-time work could be undertaken from 13 to 14, officials of the brassworkers' union attended the local inquiry and bitterly opposed the proposition, on the grounds that the Education Act of 1870 and the Factory and Workshop Regulations of 1867 had already done their members enough damage. On this evidence alone there can hardly be any doubt of the subcontractors' dominance in the brassworkers' union. Although underhands had been permitted to join, they were at this stage a very minor voice. That their general secretary was to be found resisting a regulation that was expected to reduce their supply and raise their price is itself an adequate indication of that minority position.

Such was the pattern of relationships and union tactics thrown up by the local industrial structure, organization, and productive techniques, in the latter half of the century.

After the 1880's, however, this pattern, in many of its aspects, began to undergo change. The system of production by which much of the work

was done by skilled craftsmen-subcontractors organizing their own small teams and introducing subdivision of labour to the extent they found profitable began, under the economic pressures of the late seventies and eighties, to give way to new methods. Falling prices and foreign competition made this hitherto profitable type of organization less attractive, and the more enterprising manufacturers began to seek cost reductions by substituting machinery for manual skill. With the greater use of machinery and power came the greater employment of semi-skilled labour, much of it employed directly by the factory-owner, and subcontracting began to decline.

These tendencies gathered pace after the late eighties, greatly stimulated in the new century by the growth in demand for those products susceptible to mechanical and mass-production methods, in particular those associated with the new vehicle and electrical industries. Factories became increasingly self-contained, more and more processes being carried out under the one roof. This brought an increase in the size of the typical unit, and consequent pressure on the smaller men. The gradual elimination of subcontracting continued as more and more responsibilities of supervision and technical co-ordination were forced upon employers. Increasingly the labour of semi-skilled machine operatives had to be co-ordinated with the work of skilled tool makers and machine-setters, while heavier overhead expenses compelled the employer to interest himself more directly in factory discipline and the observance of regular hours. The old-type craftsman was thus either disappearing altogether, and giving way to skilled engineers and semi-skilled 'hands', or was losing his status as a small master and becoming a disciplined wage-earner of the conventional kind.

These tendencies introduced seeds of profound change in the pattern and temper of industrial relationships. What had once been a graded hierarchy of economic classes, leading through from underhand to small master and thence to employer and factor, with no unbridgeable gulf anywhere in its length, was now tending to become polarized out into wage-earners and employers. It was becoming far easier to think in terms of 'the two sides' of industry, and to have a fairly clear picture of what each side represented. It was possible, without doing too much violence to the facts, to think of the 'employing' function as moving into the hands of the higher social groups, and of the 'wage-earning' function as becoming the only one—now that the craftsman was losing his status as a master and profit-earner—performed by the lower social groups. The lower social groups, less and less subject to this division into profit-earning craftsmen and wage-earning underhands, were beginning to come together in a new solidarity against an emergent common enemy. The evils of the industrial scene

which it had previously been easiest to identify as the product of individual selfishness it was now becoming easier to attribute to a social and economic class.

Such changes reacted in turn upon labour organization, upon union policy, and in certain cases upon the balance of power within unions.

In some instances (e.g. small-arms manufacture, where these changes came appreciably earlier than in the other staple trades and were almost complete by the early seventies) the new division of the labour force into skilled engineers on the one hand, and semi-skilled machine operators on the other, caused the old craft unions to decline and finally disappear altogether. The skilled engineers moved into the Amalgamated Society of Engineers or some other engineering union, while the semi-skilled had to wait until the few years preceding 1914 before they were scooped up by such organizations as the Workers' Union and the Amalgamated Union of Gasworkers, Brickmakers, and General Labourers. Thus in place of a craftsmen's union which attempted 'vertical' combination of the different levels of skill within the trade one finds separate groupings at the two extremes of skill reaching out to ally themselves in 'horizontal' combination with corresponding levels and types of skill elsewhere.

The consequences were that at both levels the claims, objectives, and aspirations of the two groupings were not necessarily related to the problems and potentialities of (in this case) the small-arms trade. In the case of the gunmaking engineers, policies were likely to be the outcome of decisions and attitudes shaped collectively by the engineers of the whole district, rather than by individuals concerned solely with, and moulding their attitudes in the light of, the small-arms trade alone. With regard to the semi-skilled operatives, policies likewise were shaped largely by considerations having little relevance for the small-arms trade in particular. They were affected to a considerable extent indeed by their leaders' estimates as to what the public conscience could be persuaded was the minimum tolerable standard of living.

In the brass trades, however, where a considerable degree of the old craft element remained, the craft union survived, with the interesting result that as subcontracting declined the union passed through a crisis arising from internal power struggles. The increase in the number of wage-earning craftsmen employed directly by the firm, and their increased weight in the union, began to bring about a shift of power within what had hitherto been essentially a subcontractors' organization.

The rising power of the new pressure group began to exert discernible effects on union attitudes as early as the eighties, but it was not until the nineties that the changing balance of power began to reflect itself in major policy. They were by then powerful enough to compel the union to initiate

a Minimum Wage campaign, in which the minimum was claimed not only for the wage-earning craftsmen directly employed but also for all underhands employed by subcontractors. Subcontractors were to receive union support for only such increases in piece-prices as would enable them to pay the minimum to their underhands, and be left with profits equivalent to the minimum for themselves. Subcontractors who, though left with an amount equal to or above the minimum after paying the minimum to their underhands, had nevertheless been obliged to sacrifice some of their profits to do so, received no support from the union in any further claims they made upon the employers to maintain their incomes.

The balance of forces had thus shifted substantially. Subcontractors earning less than the minimum were allied with underhands and wage-earning craftsmen against the employers and subcontractors earning more than the minimum. The initiation of this campaign was at once followed by an exodus from the union of the better-off subcontractors, and union membership slumped in the space of twelve months by nearly 30 per cent.

Not the least interesting aspect of this shift of power within the union is the way in which the old leaders survived it by showing themselves prepared—once the new balance of forces had become clear—to follow their new masters. The wage-earning craftsmen and underhands did not so much take over the union machinery as compel the old leaders to manipulate it to new ends.

The change of policy was not confined to the adoption of the Minimum Wage campaign. It emerged also in the changed union attitude towards the influx of boy labour into the trade. Along with the Minimum Wage campaign went a demand for the limitation of juvenile workers—a step which could only harm the subcontractors if it succeeded in reducing the abundance of underhand labour and pushing up its price, as of course it was designed to do.

By the early years of the new century the shift of power was complete. The union was now condemning unequivocally the practice of subcontracting. The general secretary, who in 1875 was bitterly opposing a Factory Act regulation which threatened to raise the price of underhand labour, was in 1909 inveighing equally passionately at public meetings against the whole system of subcontracting, with its 'exploitation of boy labour', and describing it as 'the curse of the trade'. Clearly the brassworkers' union had travelled a long way.

These shifts of power and the resulting changes of policy did not, however, affect the general temper and outlook towards the employers until the few years immediately preceding 1914. Probably an important factor here was the continued ascendancy of many of the old leaders; at any rate

the attitudes and philosophy of W. J. Davis continued to characterize the field. In 1911 he was still pressing his views.

> Buyers of brass and other metal commodities know no code of honour as to legitimate profits in a growing competitive age. They will take the lowest quotation, and if [it] . . . yields a profit to manufacturers and a living wage to the workers they are just as satisfied as though the price paid ruined the contractor and sweated the workers' wages. This means that the only protection a purchaser has is to put into competition two or more manufacturers, and . . . give his orders to those who do the work at the cheapest rate. This must, perforce, in a well-ordered society, lead to a working agreement between employers and workmen, and the sooner this is commenced in serious earnest the speedier will be the realization of a profitable industry to both capital and labour.

The figure of Davis has value as a symbol as well as for the intrinsic interest of his life. In his trade-union career, his philosophy of industrial relationships, his strategy and tactics, his transition from a leader of profit-earning subcontractors to a leader of disciplined factory wage-labour, he symbolized not only the pattern of industrial relations in Birmingham in the latter half of the nineteenth century but also the changes which began to overtake it in the years leading up to 1914.

RETAILING COSTS: SOME COMMENTS ON THE CENSUS OF DISTRIBUTION, 1950

By S. POLLARD and J. D. HUGHES

IT may be too early to judge whether the long-awaited Census of Distribution, the main volumes of which have now appeared,[1] will result in a greater interest among economists in the neglected field of retail distribution. Hitherto, the lack of reliable data (even the recent estimates of such basic data as the total number of shops were very wide of the mark[2]) might have excused such neglect. The information assembled in the Census[3] should go far to remove this disability at least.

Such academic and popular discussion as there was has been mainly concerned with the reputedly high costs of distribution; this paper is a contribution to the discussion of distributive costs in the light of the returns now available. Its object is, in particular, to isolate the factors which make for efficiency or inefficiency of distribution, as measured by the value of goods sold over the counter.

The most interesting details of costs given in the published returns are those relating to the labour factor, and in Section I of this paper the labour 'productivity' in shops of different sizes and trades is examined and labour costs are related to total gross margins. Section II deals with labour

[1] *Census of Distribution and Other Services, 1950*: Retail Trade, Short Report, H.M.S.O., 1952; Vol. 1. Retail and Service Trades, Area Tables, H.M.S.O., 1953; Vol. 2. Retail and Service Trades, General Tables, H.M.S.O., 1954. A third volume to cover the wholesale trades is in preparation.

[2] Most figures were greatly exaggerated:

	(*No. of shops in '000*)			
Estimate / *Area* / *Year*	*Census* / *G.B.* / *1950*	*Levy* / *U.K.* / *1942*	*Smith* / *G.B.* / *1939*	*Jefferys* / *U.K.* / *1938*
All Retail (including motor trade) . .	542	750–1,000	*c.* 750	750
Grocery and Food	256	..	..	307–63
Clothing	89	..	..	109–31
Confectioners, Tobacconists, Newsagents	66	..	..	95–113

The Census, it is true, covers only 92 per cent. of shops and omits Northern Ireland, but it includes stalls, barrows, &c., apparently excluded from other estimates. *Sources*: J. B. Jefferys, *The Distribution of Consumer Goods*, N.I.E.S.R., Study IX, 1950, pp. 120–3; Henry Smith, *Retail Distribution*, 2nd ed., 1948, p. 35; Hermann Levy, *The Shops of Britain*, 1948, pp. 32–33 (estimates based largely on official sources). Most estimates proceeded from the last known figures, those of the 1921 Census, on the assumption of a substantial increase in numbers of shops since. Professor Ford was almost alone in suggesting that the numbers of shops per head, at least, might be declining, cf. *Economic Journal*, 1935.

[3] For qualifications and limitations of the Census returns, see the Introductions and Appendices of the three volumes quoted.

productivity in the shops of the larger towns. Section III contains a discussion of the costs involved in holding stocks. Some conclusions and suggestions for any future Census are summarized in the final section.[1]

I

The Census material provides us with two ways of showing 'labour efficiency', or 'productivity', in shops of different sizes; firstly, by comparing wages and salaries with sales turnover; secondly, by comparing the number of workers employed with sales turnover.[2] The first of these ways of showing 'labour efficiency' is subject to serious limitations since the Census figures exclude the income of proprietors. In Table 28 of Vol. 2 of the Census, wages and salaries are given as a percentage of sales, but since the table covers only shops without working proprietors or unpaid helpers it embraces slightly less than a quarter of all shops and about 55 per cent. of all sales. In particular, therefore, it excludes most of the small shops. It is, moreover, calculated, not on the basis of the size of 'establishments' (i.e. shops), but of 'organizations' (i.e. firms).[3] Since a chain of small stores forming one 'organization' will fall into the same size category as a big store, Table 28 does not provide us with an accurate indication of the variation of labour efficiency with size of shop. Nevertheless, to give an indication of the results obtained by using this approach, we give the results for 'all retail trades', graded by size of 'organization' in Table I, cols. D and E.

The second approach, comparing the number of workers employed with sales turnover, is more promising. By combining the information given in a number of tables in the Census it is possible to calculate the sales per person engaged in 'establishments' graded by size, introducing the further refinement of counting part-time labour as equivalent to one-half of full time.[4] This approach has the advantage of covering all shops and the whole labour force in retail trade. Our method still does not allow for

[1] All our statistics are derived from the three volumes of the Census of Distribution published. They relate to one year, 1950, and the population figures are drawn from the same sources.

[2] Neither method can measure the 'quality' of service, though the rate of wages may offer an approach to such measurement. This question of 'quality' is provisionally ignored, in order to introduce measurement into the analysis.

[3] For explanation of terms see the Introductory Notes prefacing the Census volumes. 'Establishment' as used in the Census of Distribution means a separate place of business engaged in retailing goods (we have frequently used the word 'shop' in this article with the same meaning), while 'organization' means the undertaking operating one or more establishments.

[4] This refinement makes a large difference in a number of retail trades. While the sales per person employed are raised by it by 12·5 per cent. in all retail trade, among confectioners-tobacconists-newsagents the change amounts to 31 per cent. and among booksellers and stationers to 21 per cent. By contrast, the sales in department stores are only increased by 4 per cent. per person engaged.

TABLE I

Labour Efficiency in Retail Trade, by Size of Shop, Measured by Turnover

(£)

Group no.	*Size of shop; turnover in £ per annum*	*Average size within group (£ p.a.)* A	*Sales per person employed (£ p.a.)** B	*Efficiency index (col. B1 = 100†* C	*Salary and wage % of turnover* D	*Index (D9 = 100)* E
1	Under 1,000	488	489	100	19·9	188
2	1,000 and under 2,500	1,736	1,143	234	14·0	132
3	2,500 „ „ 5,000	3,674	1,696	347	12·6	119
4	5,000 „ „ 10,000	7,103	2,213	453	13·4	126
5	10,000 „ „ 25,000	15,005	2,794	571	12·8	121
6	25,000 „ „ 50,000	33,844	3,177	650	12·3	116
7	50,000 „ „ 100,000	67,888	3,158	646	11·3	107
8	100,000 „ „ 250,000	146,000	2,944	602	11·1	105
9	250,000 and over	619,044	2,805	574	10·6	100
	All Shops	9,268	2,415	494	11·0	104

* Part-time employment counted as = half full-time.

† Covers about one-quarter of all shops, 55 per cent. of sales. Grading of size is for 'organizations' not 'establishments' (for explanation see text).

different grades of labour,[1] but this is likely to be of minor importance as long as we limit ourselves to comparisons within trades. Where, later on, we do make comparisons between trades, we have drawn attention to the difference in wage and salary levels. Accordingly, we have taken sales per annum per person employed as our measure of 'labour efficiency'.[2]

Table I illustrates the relation between 'labour efficiency', as defined, and the size of shops measured in annual turnover. It shows striking variations in 'labour efficiency', of the order of more than six to one, between different grades of size.[3] The low figure for sales per person employed in shops with a turnover of under £1,000 (group 1) may actually be somewhat inflated, since goods withdrawn from stock for the proprietor's own use were (according to the Census form's instructions[4]) to be included in sales at full retail value. However, since the smallest size

[1] How striking and important these differences in the composition of the labour force can be may be illustrated from the 'furniture trades' and 'variety stores'. In the former, 62 per cent. of full-time employees are adult males and only 5 per cent. females under 18 years of age; in the latter about 15 per cent. are adult males and 20 per cent. females under 18.

[2] It may be said that 'intensity of labour utilization', or 'productivity', would more accurately describe what we are measuring. However, we feel that these terms are inappropriate, and for want of a better term have chosen 'labour efficiency' to denote the value of sales per assistant, which approximates to the reciprocal of the labour cost of distribution.

[3] There are necessarily wide variations between individual shops in each size group. The dispersion indicated in the Census, Vol. 2, Table 31, would, however, not seem to invalidate the conclusions drawn here.

[4] Census of Distribution, Retail Trade, Questionnaire Form 2, Section 6a.

group of shops accounted for less than 1 per cent. of total trade, the qualification is not an important one. The lesser variation in col. D as compared with col. B seems to stem from the fact that col. D, which, as we mentioned above, refers to only a quarter of the total number of shops, excludes the shops with a working proprietor which are most likely to be inefficient, and grades by 'organization'. The trends of both cols. D and B show, however, the same direction.

The importance of the considerable increase in 'labour efficiency' with size (of turnover) appears clearly when the effects of a marginal increase in turnover are considered. For the first six size groups:

From 1 to 2, a 10%	increase	in size	leads	to a	5·2%	increase	in	'labour	efficiency'.
„ 2 „ 3,	„	„	„	„	„	4·3	„	„	„	„
„ 3 „ 4,	„	„	„	„	„	3·3	„	„	„	„
„ 4 „ 5,	„	„	„	„	„	2·3	„	„	„	„
„ 5 „ 6,	„	„	„	„	„	1·1	„	„	„	„

The highest 'labour efficiency' is reached in shops having a turnover of between £25,000 and £100,000 per annum (size groups 6 and 7). Beyond this level 'labour efficiency' for retail trade as a whole declines slightly. Thus, while for the large shop further expansion of turnover may be said, approximately, to involve constant labour costs, the small shop, if it can expand its turnover, is likely to benefit greatly from lower labour costs per unit of sales, which will contribute to rapidly increasing profits.[1]

The importance of the variations in 'labour efficiency' with size of shop may be brought out by stating that in the country's 'small shops' (those with sales under £5,000), more than half of the total shops in number, 24 per cent. of the labour force accounts for only 13·5 per cent. of sales; on the other hand, in shops of sales between £25,000 and £50,000, a slightly higher sales figure is achieved by only 10·6 per cent. of the labour force. The difference is greater still between the extremes.[2]

It is possible to group shops by size, not of turnover, but of the labour force employed. Table II shows these groupings for 'all retail' and for some of the major divisions of trades.[3] The main trend of rising 'labour efficiency' with growing size is still demonstrated, but the difference between the highest and the lowest is not nearly as great as in Table I (for 'all retail' it is 100:129).

The reasons for the divergent results as between Tables I and II lie in

[1] Cf. the table showing the increase in profit rates in pharmacies with increase in turnover given in H. Levy, op. cit., p. 131. On the other hand, the discussion of selling economies of the small shop, op. cit., p. 199, should be noted.

[2] The Census figure for the total number of shops has already been mentioned. The Census figures for employment are: Retail Trade, 1,811,329 employed full-time, and 453,962 employed part-time. Service Trades (which are not included in our present analysis), 605,351 full-time and 133,762 part-time workers.

[3] Correcting for part-time employment (part-time as = half full-time).

TABLE II

*Labour Efficiency in Retail Trade, by Size of Shops Measured by Employment (Shops other than Co-operative)**

(All Retail and Grocery Annual Sales per person engaged, £)

Establishments with no. of persons engaged	*All Retail*	*Grocery*
1 person engaged	1,991	2,268
2 persons ,,	2,307	2,417
3 ,, ,,	2,394	2,548
4 ,, ,,	2,430	2,648
5–6 ,, ,,	2,455	2,778
7–9 ,, ,,	2,454	2,876
10–24 ,, ,,	2,428	2,948
25–49 ,, ,,	2,508	2,821
50–99 ,, ,,	2,510	2,763
100 and over ,, ,,	2,571	2,131
All establishments	2,415	2,692

* Part-time employment counted as = half full-time.

the form of grouping adopted. The grouping by sales (as adopted in Table I) automatically pushes any efficient smaller shop (having a small labour force but a high turnover) into the upper ranges, and tends to concentrate all efficient firms in the larger size groups, and vice versa for the inefficient small shop.[1] The grouping by 'numbers employed' adopted in Table II means that all the more inefficient shops (having a large labour force but a low turnover) except the inefficient one-man shop, will be found in the higher ranges, and the very fact of high 'labour efficiency' will bring a shop into a correspondingly lower range. What Table II may be taken to indicate, to some extent, are the gains in 'labour efficiency' due to division of labour.

In Table II, the optimum sizes for 'all retail' are in the middle and upper ranges, but if we look at particular trades we find considerable variations. Thus, while in 'clothing' and 'coal' the largest size appears to be the optimum, among 'grocery' and 'furniture' the optimum sizes are to be found between 7 and 24 employees, and in 'other food' between 50 and 99. But these variations in trend in the different 'kinds of business' (i.e. trades) are seen most clearly if we return to the original grouping by size of turnover. The results for the main retail trades are set out in Table III. There are some interesting variations as between trades.

Four trades have falling labour costs throughout; 'variety stores',

[1] But the very large shop which is relatively inefficient will probably still appear in the largest size group. Hence the 'labour efficiency' in size group 9 of Table I (£2,805) comes nearer to that of Table II for establishments employing 100+ (£2,571) than does any other comparable grouping in the two tables.

'other food', 'coal', and 'confectioners, tobacconists, newsagents'. For the last named section this falling labour cost (or rising 'labour efficiency') is most marked, and has no parallel in any other trade. It appears to reflect the high sales volume of particularly favourable selling points, for the average number employed per shop even in the largest size group (with an average turnover of £154,000) is only between 11 and 12. Among other trades 'labour efficiency' rises to a maximum (in size groups 6, 7, or 8) and then declines with further expansion in turnover, but in the case of 'booksellers and stationers', 'chemists' goods', 'furniture', and 'clothing', 'labour efficiency' rises again in the largest category.[1]

In the lowest four size groups (shops with a turnover of under £10,000) the rise in 'labour efficiency', both in absolute and relative terms, is remarkably uniform among virtually all 'kinds of business' (trades),[2] despite wide differences in gross margins between them, and despite the fact that in the higher size groups there are very considerable differences in 'labour efficiency' between trades. The only exception, with efficiency considerably below the average, is 'booksellers and stationers', whose particularly low figure may be partly accounted for by the fact that 'the multiplicity of booksellers is in some part due to considerations which are not strictly economic'.[3] 'Other food' also has an efficiency lower than the rest, while 'grocery', 'coal', and 'confectioners, tobacconists, newsagents' (the trades with gross margins lower than the average) are slightly above the rest. Nevertheless, the uniformity of growth among the small shops of so many trades is remarkable and not easy to account for.

The close correspondence in 'labour efficiency' in different sections of retail trade which is found among small shops does not hold true of the larger shops (£25,000 turnover and above). These show wide variations and marked individual characteristics, both in the absolute levels of labour efficiency and in their movements with changes in size, as can be seen from Table III. The differences in the 'labour efficiency' of the various sectors of retail trade are, therefore, largely due to the differences among the shops with a larger turnover.

It is instructive to group the sections of retail trade ('kinds of business') according to the levels of their gross margins and their 'labour efficiency'. A negative correlation between them would support the view that 'labour efficiency', as used here, exerts a strong influence on total costs and gross margins. This correlation exists for the majority of trades.[4] The exceptions

[1] Since there is a correlation between size and efficiency over most of the field, there is also a broad correspondence between average size of shop and average efficiency as between trades, e.g. furniture and coal, trades with predominantly large shops, show a very high efficiency, while trades with mainly small shops, such as jewellers and booksellers, show a very low one.

[2] Cf. Diagram 1 (Appendix).

[3] Levy, op. cit., p. 111.

[4] Lack of space forbids detailed statistical illustration.

TABLE III

Retail Trade: Sales per Person Employed, by Size of Shop (Turnover p.a.) and by Trade

(Part-time employment = half full-time)
(£000 sales per person employed)

Class	Size of shops: Turnover	Grocery	Confectioners Tobacconists Newsagents	Coal	Chemists' and Photogr. Goods	Booksellers and Stationers
1	Under £1,000	0·5	0·5	0·5	0·5	0·4
2	£1,000 and under £2,500	1·3	1·2	1·3	1·1	0·8
3	2,500 „ „ 5,000	2·0	1·8	1·8	1·6	1·1
4	5,000 „ „ 10,000	2·6	2·4	2·2	2·1	1·4
5	10,000 „ „ 25,000	3·0	3·5	2·6	2·3	1·8
6	25,000 „ „ 50,000	3·3	6·3	3·3	2·3	2·3
7	50,000 „ „ 100,000	3·4	10·4	3·7	2·2	2·9
8	100,000 „ „ 250,000	3·7	13·5	4·1	2·2	2·5
9	250,000 and over	2·9		6·6	2·3	3·2
	All sizes	2·7	2·7	2·7	2·1	1·6

Size of shops Class	Jewellery, Leather, Sports Goods	Furniture	Clothing	Hardware	Other Food	Other Non-Food	Variety Stores	Department Stores
1	0·5	0·5	0·5	0·5	0·5			
2	1·0	1·1	1·1	1·1	1·1	0·8*	0·9	
3	1·5	1·6	1·6	1·6	1·6			
4	2·0	2·2	2·2	2·1	2·0	1·6*		
5	2·6	3·0	2·9	2·7	2·5	2·0	1·6*	
6	3·1	3·5	3·2	3·0	2·7	2·4		
7	3·6	3·8	3·0	3·0	2·9	2·8	2·1	2·7
8	3·5*	3·7	2·9	3·3	3·0	2·5	2·3	
9	3·5*	3·8	3·2	2·2	3·0		2·5	2·5
All sizes	2·0	3·0	2·5	2·2	2·1	1·5	2·3	2·5

* No separate figures for part-time workers available; all counted as full-time.

are department stores, clothing, and furniture, which combine higher than average 'labour efficiency' with higher than average gross margins; these are types of trade in which heavy stocks tend to raise gross margins beyond the average.[1]

If we compare, for different 'kinds of business', the 'labour efficiency' of shops of £25,000 a year turnover and above (in which the differences are brought out most sharply) with their gross margins broadly similar results appear. High labour costs are associated with high margins, and low labour costs with low margins. Again, there are some exceptions; furniture, clothing, and jewellery show high 'labour efficiency', yet charge high margins, but all three have high costs of holding stocks. On the other hand, the wide differences in 'labour efficiency' between 'furniture' (£3,670 per assistant) and 'variety stores' (£2,294 per assistant), with gross margins of about 27 per cent. for both, points to a second compensating factor, the type of labour involved. The gap between them disappears when we take into account the different wage and salary levels of the two trades, which is due to the different types of labour employed. The 60 per cent. higher 'labour efficiency' of 'furniture', is balanced by an average wage (and salary) which is 58 per cent. higher than in 'variety stores'.[2] In any case, labour costs are not the only costs of retail distribution, and there is no reason to assume that these wide divergencies in 'labour efficiency' necessarily indicate or lead to different rates of net profit in different trades. Lower efficiency in a trade may in some cases be compensated for by lower costs of holding stocks—or by high margins. It seems clear that the influence of 'labour efficiency' on gross margins, i.e. on distributive costs, is considerable.

This last point is reinforced by the fact that, in addition to the negative correlation between 'labour efficiency' and gross margins which is found in comparing different retail trades, the same negative correlation holds if we compare different sizes of shop. Unfortunately, the Census of Distribution only provides us (in Vol. 2, Table 28) with gross margins for the different size groups (by annual turnover) of 'organizations'. Thus, an exact comparison of gross margins with 'labour efficiency' in different size grades of 'establishments' is not possible. However, the main trend of 'labour efficiency' in different size grades of 'organizations' is substantially similar to that for 'establishments', and in Table IV this is compared with gross margins.

A clear picture emerges. Distributive costs, as measured by gross margins, are lowest in the middle size ranges, rise with the larger size

[1] Gross margins in some trades are inflated by services rendered. (Cf. Section III, below.)

[2] Making allowances for part-time labour.

TABLE IV

Efficiency of 'Organizations', by Size

*Size group**	*Sales per persons employed†* (£000 *p.a.*)	*Gross margins* (% *of sales*)	*Number of stock turns‡*
1	0·4	27·2	5·6
2	1·1	22·4	7·8
3	1·5	19·9	10·2
4	2·0	19·6	10·3
5	2·5	20·3	9·0
6	2·8	22·1	7·6
7	2·9	22·6	7·4
8	2·8	23·4	7·1
9	2·6	23·8	8·6

* As in Table I, but refers to 'organizations'.

† This is the median for organizations in each group, and is not corrected for part-time labour.

‡ Sales divided by end-year stocks.

ranges, but are highest in the smallest size.[1] They follow a trend similar to that of labour costs, except that gross margins begin to rise after size group 4 (£5,000 to £10,000), before 'labour efficiency' has reached its peak. This early rise in gross margins seems to be associated with the rising burden of stocks in relation to turnover. The number of stock turns (sales divided by end-year stocks) is highest in size groups 3 and 4, i.e. the proportion of stocks carried to sales is lowest, and the number of stock turns falls thereafter until the largest size group is reached.

The negative correlation between 'labour efficiency' and gross margins which has been observed in comparing both different 'kinds of business' in retail trade and different sizes of 'organization' certainly strengthens the impression that the influence of 'labour efficiency' on distributive costs can hardly be exaggerated.

II

Labour efficiency in distribution may also be examined in the light of Tables 8 to 11 of Vol. 1 of the Census, which describe in detail the distributive trade of all larger towns. In this section only the towns with populations of over 50,000 are examined. Excluding the central London areas

[1] It is surprising, at first sight, that shops of different sizes in the same trade should show marked variations in gross margins, but Table 28 in Vol. 2 of the Census shows how general this feature is. Moreover, firms 'without working proprietors' show significantly higher gross margins than shops 'with working proprietors' (24 per cent. compared with 20 per cent.), while their average 'labour efficiency' is also appreciably higher (£2,579 p.a. compared with £2,015), a curious result which derives from the fact that the latter include most of the small firms and the former most of the large organizations. The analysis of these margins raises questions outside the scope of this paper which we hope to deal with at a later date.

(the Cities of London and Westminster and the Metropolitan Boroughs of Chelsea, Holborn, Kensington, and St. Marylebone) the Census lists 17 towns of over 250,000 inhabitants, 65 of 100,000 and under 250,000, and 100 towns of 50,000 and under 100,000, a total of 182 towns[1] representing 50·5 per cent. of the total population and 57·3 per cent. of total retail sales.

In thus grouping the towns by size[2] we find that those with over 250,000 inhabitants have a significantly higher figure of sales per head (£121 p.a. compared with £110), representing the fact that most important cities are shopping centres of their surrounding areas.[3] In commodities which are normally bought locally that difference disappears and, in fact, turns into a slight deficiency of the largest cities, probably due to the extensive poorer working-class quarters to be found in most of them:

Retail Sales of Groceries and Other Food, Per Head (£)

Towns 250,000 and over	44·23
Towns 50,000 and under 250,000 . . .	45·31
(Great Britain)	(42·60)[4]

If sales to citizens only could be isolated, the largest cities would most probably be close to the average of all towns.[5] The towns between 50,000 and 250,000 population also make substantial sales to non-citizens, but in their case this proportion is smaller, while many of them also 'lose' some of their sales, in turn, to nearby large cities. There is no appreciable difference in the average of the two groups of smaller towns.

The higher figures of sales per head in the larger cities are matched by a corresponding advantage in the size of shops (250,000 and over, £11,700 turnover p.a.; 50,000–250,000, £10,000) due, at least in part, to the large emporiums in city centres.[6] However, the labour efficiency is *not* here in

[1] Including the other Metropolitan Boroughs.

[2] In Vol. 2, Tables 7 and 10 of the Census, a different system of grading towns has been adopted, separating the Greater London area.

[3] Averages for cities at the centre of 'conurbations' are especially high: Newcastle £164, Manchester £152.

[4] The low sales in agricultural and market gardening areas reduce the figures for the country as a whole.

[5] Sheffield with £106 and Hull with £99 are examples of cities without an important 'hinterland' which are below the average of all towns.

[6]

	Greater London and towns of over 100,000 inhabitants	*Rest of Great Britain*
No. of Department Stores per 1 million inhabitants . .	15·4	7·0
No. of establishments with sales of over £50,000 per annum per 1 million inhabitants . .	267·0	146·0

any way a function of the size of shop and amounts to £2,490 per person engaged in all larger towns. This comparison thus seems to indicate that a larger turnover per head may mean larger shops, but not greater efficiency.

This conclusion is borne out by a study of all 182 towns individually. Sales per head vary astonishingly, from £211 for Worcester to £37 for Huyton with Roby, and this is to some extent reflected in the average turnover per shop (a scatter diagram relating sales per head with average shop turnover, not reproduced here, does show a broad band rising steadily to the right), a scatter diagram comparing average sales per head with labour efficiency shows, at first glance, no such correspondence.

Among the 182 towns there are, however, wide variations in social and economic character, and certain 'clusters' appearing in the latter diagram encouraged the attempt to divide the total into a number of more homogeneous groups. Such grouping is necessarily arbitrary. The majority of towns could conceivably fall into more than one group, while in view of the small number of towns the number of groups has to be kept to a minimum. In addition, proximity to larger towns may affect the shopping pattern of towns, while the large city itself is likely to appear as 'odd man out' in its group. Moreover, in the 'conurbations' the boundary lines affect rather arbitrarily the population figure assigned to each town. Six groups have been distinguished, Metropolitan, Dormitory Suburbs, Non-Industrial, New Industrial, Old Industrial, and Resorts.[1]

[1] The following is a list of the towns allocated to each group:

(*a*) Battersea, Bermondsey, Bethnal Green, Camberwell, Deptford, East Ham, Fulham, Greenwich, Hackney, Hammersmith, Islington, Lambeth, Paddington, Poplar, St. Pancras, Southwark, Stepney, Wandsworth, West Ham, and Woolwich.

(*b*) Acton, Barking, Beckenham, Bexley, Brentford-Chiswick, Bromley, Carshalton, Chislehurst-Sidcup, Coulsdon-Purley, Crosby, Croydon, Ealing, Edmonton, Enfield, Epsom-Ewell, Esher, Finchley, Hampstead, Harrow, Hayes-Harlington, Hendon, Heston-Isleworth, Hornchurch, Hornsey, Huyton-Roby, Ilford, Lewisham, Leyton, Merton-Morden, Mitcham, Orpington, Romford, Ruislip-Northwood, Solihull, Southall, Southgate, Surbiton, Sutton-Cheam, Tottenham, Twickenham, Uxbridge, Walthamstow, Wanstead-Woodford, Watford, Wembley, Willesden, Wimbledon, Wood Green.

(*c*) Aberdeen, Bedford, Cambridge, Carlisle, Edinburgh, Exeter, Gloucester, Lancaster, Lincoln, Maidstone, Peterborough, Worcester, York.

(*d*) Birmingham, Bristol, Chesterfield, Colchester, Coventry, Dagenham, Derby, Gillingham, Gosport, Grimsby, Ipswich, Leicester, Luton, Mansfield, Northampton, Norwich, Nottingham, Nuneaton, Oldbury, Oxford, Plymouth, Portsmouth, Reading, St. Helens, Scunthorpe, Slough, Southampton, Swindon, Thurrock, Walsall, West Bromwich, Wolverhampton.

(*e*) Barnsley, Barrow-in-Furness, Birkenhead, Blackburn, Bolton, Bootle, Bradford, Burnley, Bury, Cardiff, Crewe, Darlington, Dewsbury, Doncaster, Dudley, Dundee, Gateshead, Glasgow, Gt. Yarmouth, Greenock, Halifax, Huddersfield, Hull, Keighley, Leeds, Liverpool, Manchester, Merthyr Tydfil, Middlesbrough, Motherwell-Wishaw, Newcastle-on-Tyne, Newcastle-under-Lyme, Newport (Mon.), Oldham, Paisley, Preston, Rhondda, Rochdale, Rotherham, Salford, Sheffield, Smethwick, South Shields, Stockport, Stockton-on-Tees, Stoke-on-Trent, Stretford, Sunderland, Swansea, Tynemouth, Wakefield, Wallasey, Warrington, West Hartlepool, Wigan.

(*f*) Bath, Blackpool, Bournemouth, Brighton, Cheltenham, Eastbourne, Harrogate, Hastings, Hove, Poole, Southend-on-Sea, Southport, Torquay, Worthing.

The groups are remarkably homogeneous in the distributive patterns,[1] with the possible exceptions of the rather miscellaneous group (*d*). Table V shows the unweighted averages of the towns in each group, and the

TABLE V

Retail Distribution in the Large Towns. Grouped by Type of Town

Type of town	*No. of Towns* A	*Retail Sales per head* (£)			*Retail Sales per shop* (£000)			*Retail Sales per assistant* (£000)		
		Mean B	*L.Q.* C	*U.Q.* D	*Mean* E	*L.Q.* F	*U.Q.* G	*Mean* H	*L.Q.* J	*U.Q.* K
(*a*) Metropolitan .	20	104	93	117	10·3	8·9	12·5	2·78	2·70	2·88
(*b*) Dorm. Suburbs .	48	97	76	132	12·0	10·4	13·2	2·75	2·60	2·83
(*c*) Non-Industrial .	13	150	139	156	12·6	10·6	14·1	2·49	2·38	2·58
(*d*) New ,, .	32	116	98	138	10·4	8·8	12·2	2·48	2·36	2·59
(*e*) Old ,, .	55	109	96	122	8·4	6·9	10·0	2·31	2·19	2·39
(*f*) Resorts . .	14	143	127	168	10·7	9·2	12·8	2·25	2·15	2·37
All towns . .	182	114			10·6			2·49		

Notes: Cols. H. J, and K (shop assistants): part-time workers counted as half.
The totals ('All Towns') in Cols. B, E, and H are unweighted averages.

lower and upper quartiles to indicate the spread, for three indices of retail distribution: (1) retail sales per head, (2) size of shop, and (3) labour efficiency. While in some groups large shops mean high efficiency, in others the opposite relation seems to hold good, and a consideration of the character of the towns in each group may point to the reasons for this seeming paradox, and for the absence of a clear relationship on a scatter diagram describing all 182 towns.

Group (*b*) consists of prosperous towns of recent growth with large middle class populations. Its shops are generally modern, relatively well adapted to their purpose, and efficiency is high throughout, the lower quartile being well above the average of all except the Metropolitan group. Nevertheless sales per head are low, destroying a seemingly obvious corollary. The reason is clearly the large amount of shopping done in the Metropolitan area, especially the City and West End districts which are excluded altogether from this discussion.

Group (*a*), the inner ring of London boroughs, presents a paradox of a different kind. Like the outer ring, it shows low sales per head (partly a sign of poverty, partly of central shopping) with high efficiency, but in this case the 'high efficiency' is less easy to explain, particularly in view of the small turnover per shop. It cannot be shown that an undue proportion of shops is in the range of medium size (which the first section of this paper

[1] Many of the irregular members of their groups are those whose inclusion was most doubtful. Groups of 'certain' members would be still more homogeneous.

has suggested is the most efficient). Nor does the big proportion of chain stores in the area seem a sufficient reason for its higher labour efficiency, for the chain stores are not noticeably more efficient than others. Equally, it would be invidious to argue that shopkeepers and their assistants work harder there than their colleagues elsewhere. It can only be suggested that the shopping facilities of such a large built-up area encourage more economical working by a more even spread of the 'load' of peak shopping hours, so that less time is wasted by assistants in empty shops, and that the higher rents obtaining in London discourage the exploitation of this favourable situation by the opening of new premises or the enlargement of existing ones. A close study of retail distribution in this 'inner ring' might be valuable for the light it could throw on the efficiency of shopkeeping.

Group (*c*) is formed of towns which have grown up, in the main, as the administrative and shopping centres of large areas. Average sales per head are high, therefore, and there are also fairly large numbers of well-to-do middle-class residents. Shops are large, particularly in the central areas, and on the average they are efficient.

Group (*d*), the least homogeneous,[1] is very near the average in all three indices. Compared with other non-Metropolitan areas its shops, though only of average size, would be more modern, and this may be related to their rather higher labour efficiency.

Group (*e*) includes the older towns, dominated by industries which were hard hit in the inter-war depressions, and with a high proportion of working-class population. Shops tend to be older and ill-adapted (except for a few main thoroughfares) and there is a large proportion of inefficient back-street shops in the working-class areas. Shops are therefore small and inefficient on the average, and although the sales per head are higher than those for groups (*a*) and (*b*), this is due to the depression of the London figures by the exclusion of the City and the West End[2] and to the recent high earnings in certain old staple industries; sales are below those of the other towns.

Group (*f*) reveals a number of apparent paradoxes. The many non-resident shoppers of the holiday seasons and the high proportion of prosperous residents ensure a very high sales figure per head. Yet shops are of no more than average size, and 'efficiency' is exceptionally low. It must be conjectured that shops are comparatively small because stalls and barrows, many of which would not operate for more than half the year, were enumerated as 'shops' in the Census, and similarly many of the

[1] It includes the newer ports, the dockyard towns, and towns in which the dominant newer industries exist side by side with older ones.

[2] Sales of 'Greater London' as a whole are the highest for any standard region, £125 per head.

summer salesmen would turn to other occupations in the winter and thus show a low *annual* 'labour efficiency' figure. Resorts would also have a heavy proportion of sales of articles on which a high margin was earned on a low turnover.

This whole analysis in terms of types of towns might be objected to on the grounds that the grouping of the towns is arbitrary and possibly incorrect. Two further points may be made in its support.

Firstly, the Census of Distribution shows a marked relation between the relative 'prosperity' of a town and the proportion between 'groceries' and 'other food' bought in its shops (called the 'prosperity ratio' in the following discussion). 'Groceries', in 1950, consisted to a large extent of rationed and subsidized goods and grocery sales per head were fairly uniform over the whole country. 'Other food' reflected more clearly the purchasing power of shoppers, though it includes butchers and off-licence sales which tend to reduce this effect.[1] This prosperity ratio is a better index than would be the ratio between luxury goods and necessities, for it eliminates the irrelevant effect of shoppers residing outside the towns in question.[2] Towns in each of the six groups fall into a narrow range of this 'prosperity ratio', thus confirming the grouping that has been made. The ratios of the aggregates within the groups are as follows:[3]

	Other food/groceries
(*b*) Dormitory suburbs	1·263
(*a*) Metropolitan	1·241
(*f*) Resorts	1·141
(*d*) New Industrial Towns . . .	1·008
(*c*) Non ,, ,, . . .	0·862
(*e*) Old ,, ,, . . .	0·800

[1] 'Other Food' includes also fishmongers, poulterers, greengrocers, fruiterers, bakers, cooked meat and delicatessen, 'other food', and general food shops (the last three of comparatively minor importance).

[2] It might be objected that the apparent higher efficiency in 'prosperous' areas merely represents the higher prices of the goods bought there, and that on a comparison of *quantities* bought the differences would disappear. Our reasons for thinking this of slight importance are: (*a*) high-class shops need not necessarily show a higher turnover per assistant; shops of this class have to allow more time for each customer and keep a large enough staff to avoid inconvenience to the customer; (*b*) neither sales per head nor sales per shop bear the close relation to sales per assistant which one would expect if quality and unit values of goods were the main factors; (*c*) 'prosperity' is here used in a special sense. It relates largely to the inter-war years when most of the shops existing in 1950 acquired their character (by building or adaptation), and as a result of which certain buying habits remain, which are shown in the 'prosperity ratio'. By 1950 there is no reason to believe that the population of e.g. Tynemouth, Barnsley, or Sheffield had lower earnings than those of Carlisle, Luton, or Edinburgh.

[3] The ratio may be applied to other comparable areas. A comparison of Standard Regions

Secondly, an attempt has been made to check our results by eliminating the groups as such altogether and considering the towns in the light of population statistics which are independent of any arbitrary classification. Neglecting the 20 Metropolitan areas and the 13 Resorts, the classification of which is the least open to doubt, the remaining 149 towns might be expected to show some correlation between the change in population in the last 30 years and the labour efficiency in distribution: the dormitory suburbs growing fastest and with a high efficiency; the new industrial and non-industrial towns growing steadily and with average efficiency in distribution; and the old industrial areas, having a stationary or declining population, the least efficient.

The scatter diagram (Diagram 2, Appendix) relating population changes since 1921 with efficiency in distribution, strongly confirms these assumptions. The band rising from left to right is too clearly defined to be fortuitous. This is the more gratifying since population changes are only an indirect index of the patterns of distributive trading suggested here; at the same time, indices of individual towns are likely to be distorted by the proximity of large cities.

Two warnings are necessary. Firstly, the diagram is not intended to illustrate a direct causal relationship between population changes and distributive efficiency (save in one aspect mentioned below). Secondly, no proof of a linear relationship is attempted, in view of the idiosyncrasies of the 149 towns. Rather, four broad groups can be observed: the old towns, declining or increasing by a maximum of about 12 per cent. in the 30 years, with efficiencies mainly between 2·00 and 2·44,[1] higher for the slightly growing than for the declining towns; the new industrial towns, with growths of between 10 per cent. and 100 per cent., with efficiencies between 2·25 and 2·60; the non-industrial towns, a compact group, growing by 10 per cent. to 35 per cent., with efficiencies of 2·30 to 2·65; and the suburbs, growing by 30 per cent. and up to 500 per cent. and over, with efficiencies of 2·40 to 3·00 and more.

It is time to draw conclusions. Study of the geographical details of the Census shows that 'labour efficiency' is not merely a function of the size of shops in each trade or in all trades, irrespective of their conditions or surroundings. The more modern shops, in the newer areas, appear more

of Great Britain shows the following ratios (the levels are lower on the whole than for the towns because of the home production of the agricultural areas):

Greater London	1·253	South-Western	0·700
South-Eastern	0·882	Scotland	0·683
Midlands	0·840	North Midlands	0·676
Southern	0·767	East and West Riding	0·626
Eastern	0·761	Northern	0·615
North-Eastern	0·712	Wales	0·533

[1] In £000s per person employed per annum.

efficient (and this is especially true for the new, chiefly residential suburbs of the largest cities). They are possibly better sited and more expertly run. London conditions, despite many adverse factors, also seem to encourage a high degree of efficiency. At the other end of the scale, the old towns, with their poorly adapted main shopping centres, and their even less efficient units abounding in the back streets, reduce the averages for the whole country. If distribution is inefficient, the culprits are less (as has been suggested)[1] those who cashed in on the growing retail expenditure of the inter-war years by adding yet another shop to the promising areas (hoping, as we are told, to raise margins by agreement and battening on a shrinking clientele per shop), but those who refused to move out of retailing when the residential areas around them became depopulated.

III

Labour efficiency, as discussed hitherto, can only be a partial measure of efficiency in distribution. A comparatively low proportion of manpower might conceivably be combined with an exceptionally wasteful use of capital. On the capital structure of the distributive industry the Census remains silent.[2] It does, however, include estimates[3] of gross margins on turnover, the most accurate measure of the costs of distributive trades. They are given in Table VI.

Jefferys's estimates for 1938 are included for comparison. They were arrived at chiefly by information from organizations rather than individuals and they show a broadly similar trend (the only exception being the group, 'confectioners, tobacconists, newsagents'), though at a somewhat higher level throughout.[4] Most of the 12 major groups of trades are made

[1] Henry Smith, op. cit., pp. 129–31. For a criticism of Smith's analysis on somewhat different grounds see Julia Hood and B. S. Yamey, 'Imperfect Competition in Retail Trades', *Economica*, 1951, and B. S. Yamey, *The Economics of Resale Price Maintenance*, 1954, p. 88, n. 2.

[2] Other sources of information are notoriously incomplete, and discussion of net profits in distribution commonly proceeds from the assumptions of theory, rather than fact, e.g. Jefferys, pp. 107–8. Compare the brave attempt of F. R. J. Jervis, 'Profits in Large Scale Retailing', *Manchester School*, xxi/2 May 1953.

[3] Gross margins were arrived at, not by direct information, but by deducting purchases during the year from sales. The gross margin may be misleading in individual cases due to changes in the composition of stocks, but is held to be fairly accurate for large numbers. The value of stocks (col. C) is based on year-end figures and in view of stocktaking and Christmas sales may not truly represent stocks for a number of trades. It should be noted that stocks at the end of the year were considerably higher than at the beginning, and the fact (whether due to price rises or other causes) points to an over-estimation of stocks compared with sales which were, of course, spread over the whole year. Information on stocks and margins derive from an incomplete coverage, particularly among the smaller shops.

[4] Other estimates of margins differ considerably, e.g. Levy, pp. 122, 153, *et passim*. Some of the difference may be accounted for by the fact that private investigations would elicit

TABLE VI

Gross Margins in Retail Distribution

	Gross margins, %		*Stocks % of turnover 1950* C
	1950 (Census) A	*1938 (Jefferys)** B	
All Trades	22·2 (21·1)	26–28	(11·8)
Grocery	15·3 (15·2)	18·7	(7·8)
Other Food	23·8 (23·8)	27·0	(3·0)
Confectioners, Tobacconists, Newsagents	12·9 (12·3)	23·5	(6·5)
Clothing	24·8 (24·1)	33·4	(19·4)
Hardware	27·3 (21·6)	34·8	(24·7)
Books, Stationery	30·0 (27·4)	35·5	(16·8)
Chemists', Photographers' Goods	31·2 (28·7)	33·0	(20·6)
Furniture	27·5 (25·0)	38·2	(20·0)
Jewellery, Leather, and Sports Goods	30·3 (24·7)	36·9	(36·4)
General	25·7 (25·7)	..	(12·1)
Coal, Builders' Materials, Corn	23·4 (23·4)	30·02†	(6·0)

* Based on Jefferys, op. cit., Table 10, pp. 67–68, supplemented by information in the rest of the volume.

† Includes less than 75 per cent. (in value) of the commodities contained in the Census group.

up of several sub-groups (not reproduced here) which show wider variations in gross margins, from 9·2 per cent. for tobacconists to 72·7 per cent. for Post Offices with minor retail sales.

For a detailed study of distributive costs, gross margins would have to be broken up into their constituent parts, but apart from labour costs the only costs on which the Census throws light are the costs of holding stock. At first sight the variation of stocks, as given in the Census (Table VI, col. C), between different trade groups might seem to account for the high margins in at least some of the groups. The correlation in the 'Jewellery, Leather, and Sports Goods' group at one end of the scale, and in the 'Grocery', 'Other Food', and 'Confectioners, Tobacconists, Newsagents' groups at the other, would appear particularly obvious.

The influence of large stocks on costs can easily be exaggerated. By itself, the cost of holding stocks, even on levels as high as one-third of annual turnover, would be compensated by an addition of about 3·33 per cent. on selling prices if capital could be obtained at 10 per cent.[1] Stocks generally tend to raise distribution costs not so much by tying up capital

information on the normal margins, whereas the Census necessarily includes the losses incurred in unsaleable stocks, spread over the whole turnover.

[1] 'The generally accepted axiom that the relative rate of stock turn, or in other words the holding cost, is the strongest factor in determining the relative . . . retail gross margins earned on different commodities . . . (although) it is not the only factor operating' (Jefferys, p. 101) seems to assume that shopkeepers cannot borrow at reasonable rates.

as by the expense of rent (and insurance) and by the risks of deterioration. Insurance would not normally amount to more than 1 per cent. or 1½ per cent. p.a. on stocks, or say to ½ per cent. on the price of goods sold if the proportion of stocks is high. Rent of storage space would be a considerable item for slow-moving bulky goods such as furniture and furnishings, but would be negligible for some of the slowest selling goods of all—jewellery and fancy goods, while it might be high for some commodities with a rapid turnover (such as coal and greengroceries). Deterioration varies with the character of goods and may well be important even in shops carrying low stocks. A substantial stock liable to deterioration would naturally exert a considerable influence on gross margins.[1]

A number of trades offer some services which tend to inflate their gross margins to some extent. The Report gives details of the firms' own estimates of the value of their 'services', and it is possible to correct the ratios of gross margins by deducting the value of services (other than the provision of meals) from both the aggregate sales and the aggregate margins, giving 'corrected' values of gross margins. The corrected figures are shown in Table VI in parentheses. The figures for stocks have been similarly corrected.

A more detailed comparison of sub-groups of trades, of which 72 are listed in the Census, shows that generally high margins and stocks or low margins and stocks (all 'corrected' for services, frequently go together). The commodities in which a high level of stocks and their deterioration appear to account for high margins are principally articles of clothing and durable household goods; those with low stocks and margins are groceries, chocolates, sweets, tobacco, and newspapers and journals. In a number of trades the rapid deterioration of even a stock of low proportions appears to be partly responsible for high margins (perishable foods, fashion goods) but there are no commodity groups combining low margins with high stocks.

While stocks, directly and indirectly, thus have a considerable influence on retailing costs, no closer numerical relationships can be established either for the broad groupings or for the more narrowly defined sub-groups. Inevitably, other factors affecting retailing costs, such as the variations in labour efficiency between commodity groups (which were examined in Section I of this paper), the services performed, the part played by wholesalers and manufacturers in distribution, advertisement, 'unit prices', rents and other overheads,[2] taken together outweigh in many cases the influence of stocks and their perishability.

[1] Deterioration itself is already discounted in the Census figures of gross margins; but high risks would require higher than average profits.

[2] Jefferys, op. cit., part i, chap. 4; Levy, op. cit., p. 154.

IV

The foregoing analysis, although limited, has served to show that no single factor can be isolated as the main cause of 'inefficiency' or 'high costs' in distribution. Variations in distributive costs arise out of a complex of factors of varying importance, which cannot be contained in one formula. This paper has attempted to draw attention to the influence of the size of shops, the 'prosperity' of the district, and the risks and costs of holding stocks, on the costs of distribution. (There are other factors, such as the level of rents and the degree of skill of labour in distribution, on which the present Census can throw little or no light.) Each of these might be subjected to separate examination with a view to a piecemeal cost reduction without fear of reducing the intangible quality of 'service'. There are those who hold that 'the test of efficiency in retailing does not lie in the level of costs',[1] on the grounds that sales statistics cannot measure all the services rendered by the shopkeeper, but they would need to prove that such services are necessarily high when sales efficiency is low, and low when sales efficiency is high, no matter from what cause. Such proof has so far not been forthcoming.

Our analysis of the first two volumes of the 1950 Census of Distribution is aimed at stimulating a wider interest in this new addition to the statistical information available on British industries, affording for the first time the opportunity of examining many features of the distributive trades, hitherto subjects of speculation only. The present analysis, even though it has only very limited objectives, has been hampered to some extent by the fact that not all the information collected has been made available. In particular, it would have been helpful to have presented Tables 27–31 of Vol. 2, showing wages, sales per employed person, and gross margins for all sub-groups of trades, in size groups of 'establishments' (shops), not of 'organizations', for in that way the efficiency of the small shop could more easily have been compared with that of the large. As it is, chains of small shops are grouped with the largest establishments (since their aggregates are decisive), and not with other shops of comparable size. A further improvement might have been made by extending the division by size groups to each of the large towns (if only for the main groups of trades) to show the location of the mammoth stores, and the distribution of the very small shops. Similarly, the level of gross margins in larger and smaller towns might have been isolated, if only for towns grouped by size as in Tables 7 and 10, Vol. 2.

It would, however, be ungenerous to cavil at minor defects when the compilers of the Census Reports have assembled an almost miraculous

[1] Basil Yamey, 'Evolution of Shopkeeping', *Lloyds Bank Review*, Jan. 1954.

mountain of information out of the handful of answers supplied by each shopkeeper. If some of our questions are still unanswered, it is mainly because the questionnaires on which the Report is based have been limited in scope.

The following extensions in future Censuses of Distribution may, perhaps, be suggested as a minimum:

(I) At present gross margins are, in effect, divided into 'labour costs' and 'the rest'.[1] As a first approximation, the division of gross margins into at least a number of major categories should be attempted, e.g. rents, rates, heating and light; taxes; transport; improvement to fixed property; advertisements; fixed interest charges (as well as changes in stocks and wages) and miscellaneous, to account for the whole gross margin on each return completed. The most difficult of these returns for the average shopkeeper, the level of stocks at the beginning and the end of the year, has already been supplied under the present Census, and the other items should present no major problems.

(II) An attempt should be made to inquire into the capital structure of the distributive industry. A comprehensive inquiry might present insurmountable obstacles, but a beginning could be made by requiring a return of the purchase price of all retail establishments which have changed hands within, say, three years of the Census period. The details of sales, costs, &c., of this limited sample could then be presented separately (while the shops included in it would also be entered in the national tables) to see how far it was representative.

(III) The vital returns of margins and stocks are at present available for firms only, and not for individual shops; this is because firms owning more than one shop were permitted to combine the purchases and stocks (among others) of their whole organization for the purpose of the Census. They should, in future, be asked to estimate these for each of their branches, to permit a tabulation showing the efficiency of the small establishment.

[1] Even if costs of holding stocks are deducted, large margins are still unaccounted for. Thus in the following trades, if we deduct from gross margins the 'corrected' wages bill (allowing for unpaid proprietors the average wage of their trade, counting part-time as half) and 10 per cent. of the value of end-year stocks, we still have:

	Gross margins	*Deduction of wages and stocks*	*Residual margin*
		(in % of sales)	
All Retail . . .	22·2	12·5	9·7
Grocery . . .	15·3	10·2	5·1
Other Food . . .	23·8	13·8	10·0
Clothing . . .	24·8	13·4	11·4
Furniture . . .	27·5	13·5	14·0
Jewellery . . .	30·3	18·5	11·8

The drawing up of forms to be filled in by 700,000 individuals is necessarily unpopular. Nevertheless, since the successful completion of the first Census of Distribution leads us to hope that it will be followed at regular intervals by others, becoming as much accepted records of economic development as the Censuses of Production, the future Censuses should be made as complete as possible.

UNIVERSITY OF SHEFFIELD.

APPENDIX

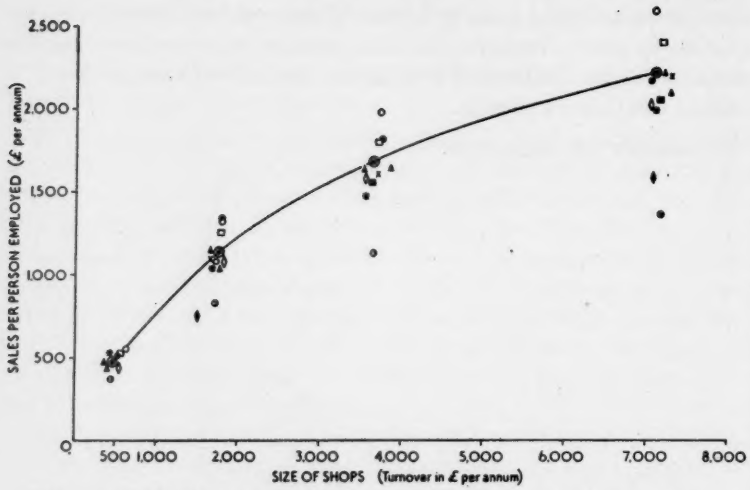


Diagram 1. *Small shops*: Increase in Sales per Person Employed with Increase in Size. (All shops —⊙—).

- ○ Grocery.
- △ Chemists' goods, &c.
- × Furniture.
- □ Confectioners, Tob., News.
- ⁕ Jewellery.
- ▲ Clothing.
- ◊ Other Food.
- ■ Hardware.
- ◑ Booksellers.
- ● Coal.
- ◆ Other Non-Food.

LABOUR EFFICIENCY
SALES PER ASSISTANT
[£1000]

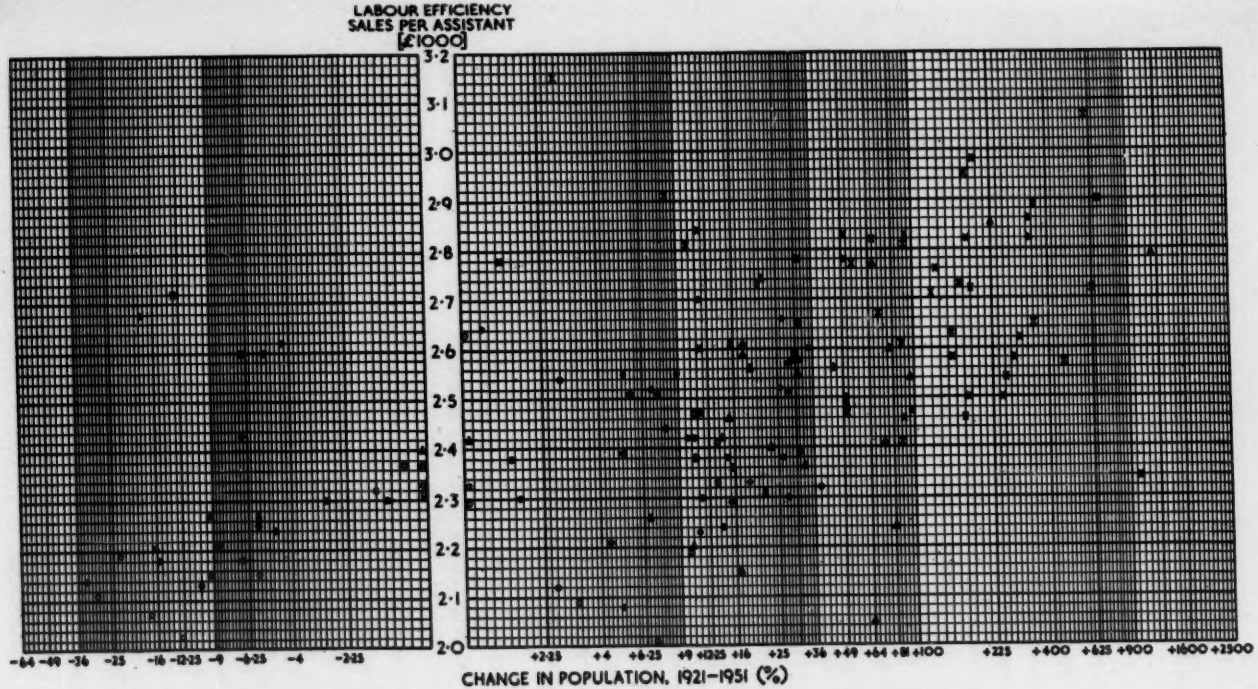


Diagram 2. Efficiency in Distribution and Population Changes in the Large Towns.

Notes. Population figures are taken from the Preliminary Censuses of England and Wales and of Scotland, 1951; they refer to the changes between 1921 and 1951; in some cases the population figures for 1951 show considerable differences from those given by the Census of Distribution for 1950, used in the remainder of this paper.

Explanation of symbols: × Dormitory and Light Industrial Suburbs; ● Non-Industrial Towns; Δ New Industrial Towns; o Old Industrial Towns.

PRICE FORMATION IN MANUFACTURING INDUSTRY AND EXCESS CAPACITY

By H. R. EDWARDS

1.0. In his recent contribution 'The Theory of Imperfect Competition Revised', Mr. R. F. Harrod is concerned primarily to refute the doctrine that conditions of 'imperfect competition with free entry' regularly give rise to the creation of excess capacity.[1] This doctrine is widely held, and has important welfare implications. Mr. Harrod's argument therefore merits the most careful attention, and the more so since it necessarily involves a reconsideration of the whole question of the theory of pricing in conditions of Imperfect Competition (in particular, manufacturing industry, with which alone we shall be concerned in this paper). I propose to examine his argument from the point of view of Mr. P. W. S. Andrews's theory set out in *Manufacturing Business*.

2.0. The starting-point of Mr. Harrod's analysis is the proposition that 'the possibility of the relatively free entry of rivals [in Imperfect Competition] is fairly widespread' (p. 145). This would seem to be in accord with the presumption of the 'accepted doctrine' of Imperfect Competition in question, and contrary to Mr. Harrod's own previous belief.[2] The precise import of the assumption of 'free entry', however, is more complex than would appear at first sight; and a proper appreciation of this point is important for the theory of the determination of prices in manufacturing industry.

The Concept of 'Free Entry'

2.1. Mr. Harrod begins with the assertion that 'the question whether free entry would often be the correct assumption' is 'a question of fact, not of analysis' (p. 143), and he therefore appeals to the 'testimony of practical persons'. Thus (pp. 143–4):

> While field inquiry fully confirms the view that imperfect competition is widely prevalent, if not omnipresent, throughout the field of manufacturing industry, a lively fear of the possibility of the incursion of new entrants, should excessive prices be charged by those already producing a certain article, is very widespread. While there are many cases of fairly securely established monopolistic (or oligopolistic) positions, in the majority of cases it seems to be assumed that sustained high prices will attract new entrants.

And further on he writes in the same sense:

> Entrepreneurs usually have the notion that there is no lack of potential rivals, should they expose themselves by charging an unduly high price.

[1] See his volume *Economic Essays* (Macmillan, 1952), pp. 140 ff. In what follows, all page references in the text refer to this volume.

[2] In his 'Doctrines of Imperfect Competition', *Q.J.E.* May 1934. See pp. 125–6 of the reprint in the present *Essays*.

Now the implications of this testimony are several. It means, for instance, precisely what it says, namely, that entrepreneurs *are aware* that high prices now will attract new entrants. This is of no little significance, since 'accepted doctrine' presumes that entrepreneurs are not so aware. In the present connexion, however, it is taken by Mr. Harrod to establish the more limited proposition: that for a wide range of manufacturing industries entry is 'free' in the sense that there are no physical, monopolistic, or financial obstacles to prevent the entry of new capacity, such as a 'corner' on essential materials or insufficient capital resources.

Now the 'accepted doctrine' in question is based on the same assumption; and so also is Mr. Andrews's theory.

2.2. But to assert that entry is *free* in this sense does not necessarily mean also that it is *easy*. The tendency has been to regard the two terms as synonymous, or rather inclusive; and Mr. Harrod is himself not blameless in this.

Consider first the case of Perfect Competition.[1] Here indeed the presumption is that there is both freedom and ease of entry. Perfect Competition envisages a standardized product, complete knowledge of all offers, and (hence) a perfectly random pairing of buyers and sellers. These conditions are realized in practice approximately where there is an organized central market for the product—as obtains, for example, in the case of most agricultural commodities. In these circumstances the new entrant—along with the established producer—is assured of the sale of his output at the price established in the market. 'Entry' requires only the necessary capital and resources to set up in production.

The case of manufacturing industry is different. Here, in general, there is no centralized marketing organization. The problems confronting the prospective new entrant, therefore, are of a different order as compared with the case of Perfect Competition. The new entrant in manufacturing industry has not only to establish his production unit; *he has also to acquire the essential 'connexion' in the market*, which is the *conditio sine qua non* of his efficient and continuing manufacture. And to appropriate this is not in general a simple matter, since it involves a permanent alteration in the established channels of trade. In this sense entry, while free, is not easy.

The initial difficulties confronting the new entrant are formidable indeed. Thus in the case of manufactured intermediate products, he must contend with the acquired 'business goodwill' of the established firms. The buyers of the product—themselves manufacturers at a higher stage of production—will tend to look first to their customary suppliers because of the confidence gained by previous custom, and it is from these firms that they

[1] On this, see my article 'Goodwill and the Normal Cost Theory of Price', *Economic Record* (Melbourne), May 1952, pp. 58–62.

will normally buy. The product the would-be new entrant would offer may be in all respects of unimpeachable quality. But so long as the price of the customer-firm's present supplier is 'right', why should the firm buy elsewhere? That product and its associated services (including regularity of the supply) have been proved on the basis of past experience; besides —as such firms are wont to affirm—'our existing suppliers look to us for orders'.[1]

In the case of consumer goods advertised and sold under the manufacturers' own brands, the difficulty is of a somewhat different character. In this situation the absolute magnitude of the advertising expenditure of each of the established firms is normally very large, and to a considerable degree independent of any one firm's actual volume of sales—being determined by such factors as the nature of the product, the concentration pattern of consumers in the total market area, and so on. Spread over the total sales of the established business, this advertising cost per unit is small; yet for the new entrant, striving to gain a foothold in the market, its incidence is severe[2]—and more so, to the extent that the existing firms may step up their own advertising expenditures to counter this new threat to their markets.

Thus in manufacturing industry, to force an entry in an established line of business is normally very difficult. Indeed so long as the prevailing price for the product is 'right', i.e. if it is equal to the average inclusive cost of producing the product in the most efficient manner, then new entry on a completely static view is virtually impossible.

The doctrine of entry appropriate to the analysis of manufacturing industry may therefore be enunciated as follows: that *if entry is in all relevant respects 'free', it may nevertheless be 'closed' by the very effective and legitimate weapon of a competitive price based on the costs of efficient production.*

This is essentially Mr. Andrews's view. Thus at the outset of his analysis of price formation (*Man. Bus.*, ch. v), after a preliminary discussion of the difficulties confronting the would-be new entrant, he concludes (p. 153):

> The general situation, then [in 'the market for the typical manufacturing business' (p. 148)] is that so long as its price is right, an established business will have a more or less clearly defined market, and will be protected from the efforts of would-be competing businesses to cut into that market.

2.3. But in the event that the price is *not* right, i.e. if it exceeds the 'cost of production' in the sense indicated, then experience suggests not merely that new competition will appear (and promptly), but also that *at least as*

[1] See Andrews, op. cit., p. 151.

[2] In these initial stages, of course, the new entrant's expenditure on advertising (including free samples, coupon offers, &c.) is properly a charge on capital—being an *investment* in establishing a market.

often as not the new competitor will succeed ultimately in appropriating a sufficient share of the custom of the market to permit efficient production.

It is this aspect of the whole question of entry on which Mr. Harrod would focus attention in the essay under consideration (pp. 143–4). Indeed it would scarcely be necessary to argue this point further, were it not for the preceding emphasis on the difficulties confronting the would-be new entrant; because in 'accepted doctrine' also, it is only too readily assumed (to use Mr. Harrod's own words, in the earlier context referred to above) that 'in spite of imperfect competition, new competitors can enter the business and find a market as easily as pre-existent sources'. This presumption, however, Mr. Harrod himself describes in that context as 'unlikely'; as indeed any experienced entrepreneur will affirm.

Two points then should be noticed. The first of these is that where the prevailing price is too high in the present sense, the new entrant is thereby enabled to quote a lower price, and a price which—if the necessary scale of production can be achieved—it will be possible to maintain in the long run and still earn a 'normal' profit. It is this circumstance *par excellence* which affords the new entrant the initial access to the market, which otherwise is virtually closed to him.

This is because the new low price provides a compelling focal point for his initial sales campaign. At the lower price the businessmen-buyers of intermediate products, who normally will not purchase from other than their customary suppliers, must needs consider the new product. In the event that the price has been maintained high, gratuitously, by otherwise efficient producers, it will very soon be reduced to match the price of the new entrant. But the resentment of the buyer-firms at the high price now revealed to have been not warranted by costs, provides a reservoir of ill-will which, properly exploited, will ensure the new entrant access to the market.

If the price is high because the established producers are inefficient, the success of the new entrant is the more likely. In this event the existing producers will tend to maintain their prices for a longer interval, since according to their lights they are charging a price 'equal to the cost of production', and are likely therefore to adopt the attitude that if they 'give the new entrant enough rope, he will hang himself'. The latter will nevertheless exploit the price differential to the full in his sales campaign, and the longer it is maintained the larger and the better-secured will be the footing he gains in the market. Sooner or later, however, the existing producers will be constrained to match the lower price. But for the least efficient of them, this will probably mean a quick demise; and the new entrant will stand an equal chance of capturing the custom of these firms. His place in the market will then be secure.

will normally buy. The product the would-be new entrant would offer may be in all respects of unimpeachable quality. But so long as the price of the customer-firm's present supplier is 'right', why should the firm buy elsewhere? That product and its associated services (including regularity of the supply) have been proved on the basis of past experience; besides —as such firms are wont to affirm—'our existing suppliers look to us for orders'.[1]

In the case of consumer goods advertised and sold under the manufacturers' own brands, the difficulty is of a somewhat different character. In this situation the absolute magnitude of the advertising expenditure of each of the established firms is normally very large, and to a considerable degree independent of any one firm's actual volume of sales—being determined by such factors as the nature of the product, the concentration pattern of consumers in the total market area, and so on. Spread over the total sales of the established business, this advertising cost per unit is small; yet for the new entrant, striving to gain a foothold in the market, its incidence is severe[2]—and more so, to the extent that the existing firms may step up their own advertising expenditures to counter this new threat to their markets.

Thus in manufacturing industry, to force an entry in an established line of business is normally very difficult. Indeed so long as the prevailing price for the product is 'right', i.e. if it is equal to the average inclusive cost of producing the product in the most efficient manner, then new entry on a completely static view is virtually impossible.

The doctrine of entry appropriate to the analysis of manufacturing industry may therefore be enunciated as follows: that *if entry is in all relevant respects 'free', it may nevertheless be 'closed' by the very effective and legitimate weapon of a competitive price based on the costs of efficient production.*

This is essentially Mr. Andrews's view. Thus at the outset of his analysis of price formation (*Man. Bus.*, ch. v), after a preliminary discussion of the difficulties confronting the would-be new entrant, he concludes (p. 153):

> The general situation, then [in 'the market for the typical manufacturing business' (p. 148)] is that so long as its price is right, an established business will have a more or less clearly defined market, and will be protected from the efforts of would-be competing businesses to cut into that market.

2.3. But in the event that the price is *not* right, i.e. if it exceeds the 'cost of production' in the sense indicated, then experience suggests not merely that new competition will appear (and promptly), but also that *at least as*

[1] See Andrews, op. cit., p. 151.

[2] In these initial stages, of course, the new entrant's expenditure on advertising (including free samples, coupon offers, &c.) is properly a charge on capital—being an *investment* in establishing a market.

often as not the new competitor will succeed ultimately in appropriating a sufficient share of the custom of the market to permit efficient production.

It is this aspect of the whole question of entry on which Mr. Harrod would focus attention in the essay under consideration (pp. 143–4). Indeed it would scarcely be necessary to argue this point further, were it not for the preceding emphasis on the difficulties confronting the would-be new entrant; because in 'accepted doctrine' also, it is only too readily assumed (to use Mr. Harrod's own words, in the earlier context referred to above) that 'in spite of imperfect competition, new competitors can enter the business and find a market as easily as pre-existent sources'. This presumption, however, Mr. Harrod himself describes in that context as 'unlikely'; as indeed any experienced entrepreneur will affirm.

Two points then should be noticed. The first of these is that where the prevailing price is too high in the present sense, the new entrant is thereby enabled to quote a lower price, and a price which—if the necessary scale of production can be achieved—it will be possible to maintain in the long run and still earn a 'normal' profit. It is this circumstance *par excellence* which affords the new entrant the initial access to the market, which otherwise is virtually closed to him.

This is because the new low price provides a compelling focal point for his initial sales campaign. At the lower price the businessmen-buyers of intermediate products, who normally will not purchase from other than their customary suppliers, must needs consider the new product. In the event that the price has been maintained high, gratuitously, by otherwise efficient producers, it will very soon be reduced to match the price of the new entrant. But the resentment of the buyer-firms at the high price now revealed to have been not warranted by costs, provides a reservoir of ill-will which, properly exploited, will ensure the new entrant access to the market.

If the price is high because the established producers are inefficient, the success of the new entrant is the more likely. In this event the existing producers will tend to maintain their prices for a longer interval, since according to their lights they are charging a price 'equal to the cost of production', and are likely therefore to adopt the attitude that if they 'give the new entrant enough rope, he will hang himself'. The latter will nevertheless exploit the price differential to the full in his sales campaign, and the longer it is maintained the larger and the better-secured will be the footing he gains in the market. Sooner or later, however, the existing producers will be constrained to match the lower price. But for the least efficient of them, this will probably mean a quick demise; and the new entrant will stand an equal chance of capturing the custom of these firms. His place in the market will then be secure.

Consider next the case of consumer-goods referred to in the preceding section. The ordinary consumer's attachment to the branded product of a particular manufacturer is much less stable and exclusive than the businessman-customer's preference in the case of intermediate products. As compared with the latter, the consumer has little to lose by experimenting with different products in response to lower prices and the promptings of persuasive advertising. The majority of consumers, therefore, when the product of the new entrant is widely advertised at the new low price, will at least give the new product a try. After some interval, the price will be matched; but some of these consumers, because of 'magic' effects in using the new product the first time, or perhaps for no better reason than the normal, but perverse (from the established sellers' points of view!) human desire for a change, will be won over more or less permanently to its use.

A price which is too high in relation to the cost of (efficient) production is thus a powerful stimulus to new competition because it facilitates the otherwise formidable task of forcing an entry in the market. Subsequently chance and the normal turnover of the buying population will play their part in increasing the new entrant's share of the market; in practice, conditions never remain static for long. For example, in the case of intermediate products, dynamic fluctuations in the relative positions of the consumer-firms in their own markets are just as likely to favour the new entrant as the established producer; the new entrant may grow with the success of his customers at the higher stage of production. Or the new buyer appointed in a consumer-firm may favour the new entrant. Or again, in times of seasonal pressure of demand on existing capacity, the new entrant by prompt delivery and other services may add permanently to his regular clientele. And so on. In the consumer-goods case, the position is rather that, granted the initial access and given sustained advertising and maintained quality, the name of the product will in time become as familiar as those of its rivals; and when this has been achieved it will stand an equal chance with the latter in appropriating the custom of the new buyer, of being named as the alternative choice of a consumer who finds his customary brand temporarily out of stock, and so on.

It is not necessary to prolong this discussion of the variety of business stratagems and processes by which in time, and granted the initial access, the essential 'connexion' is built up. It is clear however that time, and the willingness and ability to bear losses in the initial stages, are of the essence of the matter. And this brings us to the second of the two points to be noted in this connexion.

This is, as Mr. Andrews has emphasized, that the 'new entrant' is often, indeed usually, an established business in an adjacent line of production

where the technical processes of production are similar.[1] Such a firm, limited in the markets for its existing products (in terms of the 'goodwill' it commands), is always looking around for profitable ways of expanding its operations. So long then as one thinks in terms of the brand-new business, as is customary, one may well doubt whether new entry can be at all effective in view of the inevitable losses which the new entrant must incur in the period before the necessary minimum rate of sales (production) is achieved. But—granted the profitability of its primary markets—the new entrant which is an established firm in an adjacent line of production will be able to meet these losses, and to persist in its endeavour so long as there is any reasonable prospect that time will justify its efforts, for a much longer period than the limited resources of the brand-new firm would permit. (One offset to these losses will be the virtual tax rebate it will gain, in so far as the losses so incurred will be consolidated with—and will thus reduce—the taxable profits on its customary lines.)

We conclude that in the event that the prevailing price is too high, new competition will appear, and the high price will facilitate its entry into the market. There have always, of course, been cases where the new entrant has failed because the luck of the game has gone consistently against him; contrariwise, there are cases where, by a happy combination of circumstances, the new entrant has risen quickly to a position of dominance in the industry.[2] Reality is very complex. But *a priori* there would seem no reason why we should not assume that, at least as often as not, the resolute new entrant backed by the necessary financial resources will ultimately succeed in carving out a sufficient share of the market to enable efficient production.

A well-documented example[3] is provided by the rapid emergence and subsequent growth of new cigarette brands in the United States' cigarette industry, following the gratuitous price increase in June 1931 by the four major cigarette companies, viz. American Ltd., Reynolds, Liggett & Myers, and Lorillard (manufacturing the so-called 'standard brands' of cigarettes, 'Lucky Strikes', 'Camels', 'Chesterfields', and 'Old Golds' respectively).

Reynolds, the recognized price leader, initiated the price increase, which was of the order of 7 per cent. (an increase of 1c. in the retail price of 15c. per packet of the standard brands). At the same time sales—and the prices of tobacco leaf—were declining with the onset of the depression. The resultant 'wide margin between leaf

[1] In some cases this must be broadly interpreted indeed. Thus in the American soft drinks industry, the development of new processes for marketing 'pop' in tin cans has led to new competition from brewers, fruit juice canners, and even the can manufacturers themselves! (*Economist*, 26 June 1954, p. 1059.)

[2] As, for example, W. H. Lever (first Lord Leverhulme) in the British soap industry in the latter part of the nineteenth century. Lever's success was attributable to an innovation in marketing method, exploited by his own commercial genius. See Mr. Charles Wilson's excellent *History of Unilever* (Cassell, London, 1954), vol. i, ch. iii et seq.

[3] See R. B. Tennant, *The American Cigarette Industry* (Yale, 1950), pp. 88 ff.; and W. H. Nichols, *Price Policies in the Cigarette Industry* (Nashville, 1951), pp. 85 ff.

costs and selling prices' invited new competition; and the so-called 'economy brands' —a 'satisfactory cigarette' but of somewhat different quality from the standard brands and priced (retail) at 10c. per packet—were developed 'and prospered greatly' (Tennant, p. 89).

In 1930 the combined sales of the four major companies amounted to 97·9 per cent. of total sales. (In the previous 20 years, i.e. since the dissolution of the Trust in 1911, these companies had together consistently produced more than 95 per cent. of the total.) Following the price increase, their combined share fell rapidly, amounting to 86·6 per cent. of total sales in 1932, 82·5 per cent. in 1934, 79·7 per cent. in 1936, 73·7 per cent. in 1939—and this notwithstanding an initial cut-back in prices by the major companies in January 1933 to a level below that ruling before the price increase. By that time, however, 'a certain proportion of the smoking public which had adopted the cheaper cigarettes appears to have developed a liking for them' (Tennant, p. 90), i.e. the access to the market had been achieved.

The new brands were introduced in the main by existing old-established companies. There was, for example, the Brown and Williamson Tobacco Company, earlier an independent plug tobacco manufacturer, but by this time a wholly owned subsidiary of the British-American Tobacco Co. and the 'instrument' of British interests in the American market (Tennant, p. 67). This company reduced the price of its brand 'Wings' from 15c. to the 'economy' 10c., and succeeded in increasing its sales from 0·2 per cent. of total sales in 1931 to 6·9 per cent. in 1932 and 10·6 per cent. in 1939. (Thus, from 1932 on, it displaced Lorillard as the fourth largest producer.) Stephano Brothers, another of the 'new' competitors, was 'one of the old independent Turkish producers' (Tennant, p. 89). This firm introduced 'Marvels' and 'Tipt' cigarettes, and by 1939 had appropriated 3·3 per cent. of total sales.

In view of these figures, there would seem no reasonable doubt that most of the new brands attained a sufficient volume of sales to ensure efficient production. It is probably correct to say that, in the case of the standard brands, the cost of marketing (and particularly the cost of national advertising) was the dominating consideration in determining the minimum rate of sales necessary for profitable production, rather than considerations of production cost as such. In the present instance, the 'new' competitors overcame this obstacle by the stratagem of side-stepping it, and relying in the main on a maximum price differential (and the prevailing depression of incomes) to effect the same object.

2.3.1. When all this has been said, however, it is necessary to add that in the context of the controversy about excess capacity, the foregoing result must yet be regarded as provisional at this stage. For whether entry is difficult or otherwise, the final result of the whole system of 'accepted doctrine' is that the new entrant—along with each of the established producers—is necessarily limited to an uneconomic scale of production; this is the excess capacity doctrine in question. This doctrine will be assessed, finally, in § 5.0 below, where it will be argued, in agreement with Mr. Harrod, that it is without foundation.

2.4. We are now in a position to sum up this discussion of the assumption of 'free entry'. The exponents of the 'accepted doctrine' of Imperfect Competition would agree with Mr. Harrod and Mr. Andrews that for a wide range of manufacturing industries entry is 'free' in the sense of § 2.1 above. In 'accepted doctrine', however, it is assumed that entry is also easy—quite inconsistently, in view of its primary emphasis on

the particularity of each firm's market.[1] Mr. Harrod, in repudiating (p. 143) his own early misgivings about this 'unlikely' combination of assumptions, in effect subscribes to the same view; though not unambiguously in that, in the passages quoted above, the possibility of new entry is asserted in the context of 'sustained high prices'. This brings us to Mr. Andrews's position.

It is, first, that the entry of new competition in manufacturing industry is normally very difficult; but, second, that notwithstanding this, in the event that the price is too high new competition will come in, and—at least as often as not—will succeed in time in appropriating a sufficient share of the market.

This doctrine of entry is basic to Mr. Andrews's theory of price. It is granted that there is a certain dialectical tension as between its two parts; nevertheless it is, I think, in accord with experience.

It should be noted that by 'normally' in the first part of this statement, it is to be understood that conditions are 'settled', or established, and that the price is 'right'. The possibility of the latter is conditioned by the former. In the early stages of the development of a new industry, entry very often is indeed easy, not only (or even mainly) because the design of the product and the technique of production are primitive by the standard of later developments, but because as the product 'takes on' the aggregate demand expands rapidly of its own momentum, so that custom is easily acquired. Hence the 'mushroom' of small-scale, high-cost producers characteristic of the early development of a new industry. In time, however, as the rate of expansion slackens and the process of production is consolidated in order to realize the full potentialities of the production function, a price equal to the cost of production in the most efficient manner—the 'right' price—is possible. It is at this stage ('settled' conditions), and if such a price prevails, that new entry is very difficult. This setting of the case should always be borne in mind.

The Formation of Prices in Manufacturing Industry

(I) Mr. Harrod v. 'Accepted Doctrine'

3.0. Now it is not difficult to see how the false juxtaposition of the assumptions of easy entry and the particularity of each firm's market emerged in the theory of Imperfect Competition. That theory was developed rapidly in the early thirties, when it was recognized that it was

[1] Mrs. Robinson is quite explicit that this is, in fact, the particular combination of assumptions envisaged in 'accepted doctrine'. Thus, situations in which there are 'an excessive number of businesses with low turnover' (she writes) arise 'where the imperfection of markets [is] combined with ease of entry'. ('Imperfect Competition Revisited', *Economic Journal*, Sept. 1953, pp. 580–1.)

not true in general that the individual seller had a random relation to the custom of the market, as assumed in Perfect Competition. At that time, the neo-Marshallian Perfect Competition model had but recently attained its full development. The broad structure of that model was then taken over, far too simply, in the theory of Imperfect Competition. It was modified to take account of the uniqueness of each producer's particular market; but the assumption of easy entry—appropriate enough in the case of Perfect Competition, but entirely inappropriate in an analysis relating to manufacturing industry—was implicitly retained along with the other features of the model.[1]

And more than that. In Perfect Competition the individual producer is not at all concerned either with established competitors, or with the possibility of new competition. For the producer can sell all he wants to at the going price; and if the overall situation warrants it, new competitors will come in anyway.

Now this sort of behaviour of the individual producer in Perfect Competition was also carried over into the 'accepted doctrine' of Imperfect Competition in question: explicitly as regards established competitors,[2] and implicitly in the case of potential competitors.[3] This behaviour, however (like the presumption of easy entry), is entirely inappropriate in a model relating to manufacturing industry.

Now the burden of Mr. Harrod's argument is to deny the implicit assumption of 'accepted doctrine' that the entrepreneur is not aware of, or ignores, the effect of his present policy on new competition. Building (in effect) on the foundation of the explicit 'large group' assumptions of Professor Chamberlin,[4] Mr. Harrod assumes *in addition*—and this is not excluded by those assumptions, though it might well be argued that in such case it is not very apposite—that each producer (or, rather, the 'representative' producer) is aware that 'excess' profits now are likely to affect all markets in the future by stimulating new competition. He goes on to conclude (his further argument in this connexion will be noticed in the following § 3.1) that the representative entrepreneur 'will plan to

[1] Subsequently the position has often been rationalized in terms of spatial competition (see Professor Chamberlin's 'Monopolistic Competition Revisited', *Economica*, Nov. 1951, *passim*). In a pure spatial model, ease of entry is readily combined with the uniqueness of each firm's market.

[2] By the assumptions (*a*) that there are many small producers, and (*b*) that 'consumers' preferences [are] evenly distributed among the different varieties' of the product. (E. H. Chamberlin, *Theory of Monopolistic Competition*, 4th ed., p. 83.) It will be appreciated, of course, that the excess capacity result follows *necessarily* only in this 'large group' case of Professor Chamberlin's.

[3] The possibility that the entrepreneur should have regard to potential competitors, since it does not arise in the Perfect Competition case, was—so it would seem—merely overlooked.

[4] That this is what Mr. Harrod is doing is clear from the form of his contention (p. 151) that his reasoning does not involve 'oligopoly' *in the traditional sense*.

charge a price yielding only a normal profit' (p. 151), in which event 'the feared incursion of new entrants will not take place, and each firm will be able to retain its original market' (p. 147).

In effect, then, we have two cases: (i) the 'large group' where the 'representative' entrepreneur acts with a regard to potential competition (Harrod case); and (ii) the 'large group' where the firm does not so act (Chamberlin case, the 'accepted doctrine' in question).[1] In Fig. 1, if *d* represents the firm's initial 'particular-market' demand curve, and *LRAC* its long-run average cost curve including 'normal' profits, the final equilibrium of the firm will obtain at A in case (i), and at B in case (ii).[2]

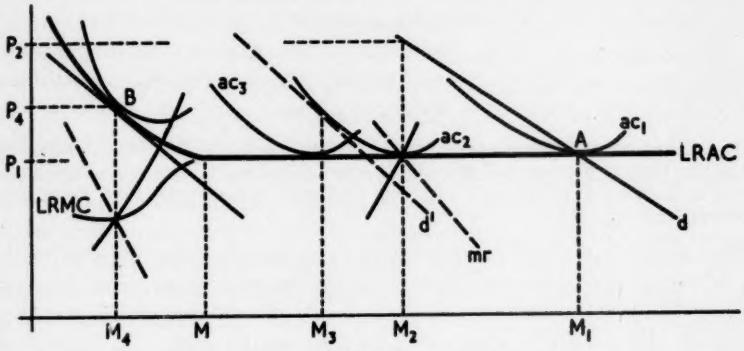


FIG. 1.

3.1. In my opinion Mr. Harrod is undoubtedly correct in his insistence, first, that established entrepreneurs are in most cases aware of the effect of high prices and profits in stimulating new competition, and second, that they will moderate their price policies accordingly.

The first part of this proposition is a question of fact, and the testimony of entrepreneurs is therefore conclusive. For my own part, certainly, I have yet to meet an entrepreneur—or even a student of elementary economics—who was not so aware. (It is, of course, an entirely different

[1] In Mr. Harrod's case the 'representative' entrepreneur (and hence all, and the group collectively) is acting oligopolistically, in the modern usage of the term. This is a matter of specific assumption; not the result of prior (structural) conditions of the case. Hence Mr. Harrod's denial that his analysis implies oligopoly.

Compare Professor Hicks's cases of 'oligopolistic expectations' and 'polypolistic expectations', in 'The Process of Imperfect Competition', *O.E.P.* Feb. 1954, p. 47.

[2] According to 'accepted doctrine', the firm would initially install the plant ac_3 (intersection of *LRMC* and the *mr* curves), and charge the price OP_2 with planned throughput OM_3. With the entry of new competitors, the *d* curve will shift to the left. With the curve at *d'*, the firm could still cover its full cost. But there can be no final equilibrium until the curve is tangential to *both* the long- and the short-run cost curves, as at *B*; for with the curve at *d'*, new firms which installed the plant ac_2 would be able to undercut the firms with plant ac_3 and so drive them out of business.

question whether, granted this awareness, entrepreneurs can judge in any particular case what constitutes a 'high' price, or otherwise. Mr. Harrod, in assuming the *LRAC* curve including 'normal' profit as known, thereby assumes that they can.[1])

As to the second part of the proposition, Mr. Harrod argues (p. 147) that the alternative course of action, i.e. charging a high price 'for the sake of a fleeting surplus profit', would involve a 'sacrifice' of 'markets' which would make the firm 'weaker' in facing 'the various contingencies of an unforeseeable future'; rather therefore 'it will . . . seek immediately to entrench itself in as large a market as it profitably can'. With some elaboration of the sense in which the firm is made 'weaker', I should be in substantial agreement with this argument.

The question inevitably arises in this context as to the proper interpretation of the basic motivating assumption, that the object of the firm is to maximize profits. Mrs. Robinson writes:

> . . . The most valid simple generalisation is that the aim of the entrepreneur is for the firm first to survive, and secondly to grow. To this end he must pursue profit, but he must avoid action which, though profitable in the present, will damage his future position. . . .[2]

This seems to me a fair statement of the businessman's approach to the profit-maximizing process. Now a regular custom, in terms of 'business goodwill', is the essential basis of the firm's efficient and continuing business, in the absence of a central marketing organization. But high prices now, which are subsequently revealed to have been not justified by costs, are well calculated to alienate that goodwill. Moreover goodwill 'snowballs'; the firm that is large and growing, and which has built up a reputation for 'quality at the price' and fair dealing is thereby enabled to appropriate more than an equal share of the new and floating demand in the market. (This, presumably, is the rationale of the oft-quoted remark by businessmen that 'The business must either grow or decline; it cannot remain still'.) Thus it is because the 'sacrifice' of 'markets' is synonymous with the loss of goodwill, and of the opportunity to extend that goodwill, that the firm is made 'weaker' by charging a high price to reap temporary surplus profits; this loss is the very negative of what is required if the firm is 'to survive and to grow'.

Normally, therefore, the firm will moderate its price policy because of a regard to the threat of new competition if the price is set too high.

This contention of Mr. Harrod's is an essential aspect of Mr. Andrews's system. As indicated in the discussion of § 2.3 above, however, the probability of the ultimate success of the new entrant if the price is set too high

[1] See Mr. Harrod's masterly essay in persuasion, pp. 155–7; and compare § 4.0 below.
[2] Op. cit., p. 582.

is not unity. This admits the possibility of alternative behaviour, as, for example, the 1931 price increase by the American cigarette manufacturers referred to above. That case, however, is indeed just the sort of exception which exemplifies the rule. Professor Tennant has rightly described it as 'a pricing error'.

Two further points should be mentioned briefly, which are not made by Mr. Harrod because of his preoccupation with the 'large group'. First, in so far as entrepreneurs *are* restrained in their price policies by their regard to the threat of potential competition, the number of *actual* competitors is irrelevant.[1] Thus the smallness of numbers in the typical situation in practice does not necessarily invalidate the theory. Indeed it may reinforce the tendency of potential competition to restrain the price policy of the entrepreneur; where the number of producers is small, then the effect of but one new competitor may be relatively large, and its approximate order of magnitude calculable.

Second, the small number of competitors having once emerged, the sort of pricing policy here envisaged will tend to maintain the position. We thus have the paradox that *the fewness of numbers—and the absence of new competitors—is often the best indication of the strength of competition* throughout the field.[2] This interpretation is often overlooked when the complaint is raised—'such-and-such an industry [the British motor-car industry in the past few decades?] is clearly not competitive, since there hasn't been a new entrant in years'. The position is paradoxical only because of the long-standing associations of 'competition' with 'large numbers', and 'fewness of numbers' (oligopoly) with a tendency to monopoly prices, profits, and outputs.[3]

(II) Mr. Andrews v. Mr. Harrod

4.0. To the extent that it envisages that the entrepreneur will have regard to the threat of new competition if the price is set too high, Mr. Harrod's 'large group' case is more closely in accord with reality than is the 'accepted doctrine' of Imperfect Competition, which followed much too simply the pattern of the Perfect Competition analysis. But in building on the basic assumptions of 'accepted doctrine' as foundation, Mr. Harrod's analysis is itself open to criticism and, in its turn, accords less closely to reality than does Mr. Andrews's system.

4.1. First, to come back to the doctrine of entry, as discussed above. The notion of 'equilibrium' requires a balance of forces with no further tendency to change. There is thus an equilibrium at A (Fig. 1), according

[1] See Andrews, op. cit., p. 153, n. 2.

[2] D. H. MacGregor, *Industrial Combination* (London, 1906), p. 6.

[3] Thus, for example, see E. H. Chamberlin, *Theory of Monopolistic Competition* (4th ed.), ch. iii, § 4, pp. 46–51.

to Mr. Harrod, because at the price OP_1 the firm is earning only a 'normal' profit on its capital and with this there will be no tendency for new entrants to come in. *But why no such tendency?*

Mr. Harrod would agree with Mr. Andrews that it is the established firm in an adjacent line of production which is the prospective new entrant *par excellence*, and such a firm (he argues) maintains 'a vigilant look-out for profitable ways of expanding its operations' (p. 144). Moreover, any firm of such a name and reputation will be seeking avenues of expansion of a more or less permanent character, and in such would deem a 'normal' return on its capital quite sufficiently 'profitable'. Surely then, if entry is easy (as Mr. Harrod in effect assumes, after the manner of 'accepted doctrine'), such firms will be very interested in lines of production where the prevailing rate of profit is in fact 'normal'? But if this is so, such a situation as that at A is not, after all, an equilibrium one; there will be a continuing tendency for new capacity to enter the line of production. It is apparent that the weakness of Mr. Harrod's argument is the too-simple assumption that entry is easy, and a function only of the level of profit, being zero when profits are 'normal'. On the contrary, as we have seen, entry is normally very difficult, and is a function of the relation of the price to the cost of efficient production. Thus while abnormal profits will always mean a price which is 'too high', the converse is not true; the price of a product may be too high (and attract new entry) because the existing producers are inefficient, albeit the price is equal to their cost of production including only a 'normal' (or even a 'subnormal') profit. It is in terms of the prospective gross profit margin, i.e. the difference between the prevailing price for a product and the entrepreneur's own estimate of the unit direct costs of production at efficient scales of production, that the prospects of successful new entry are in fact assessed.

If, then, in Fig. 1 the *LRAC* curve in fact portrays the cost of production in the most efficient manner for each scale of production, then the price OP_1 charged by the entrepreneur at A is 'right', and new entry would be very difficult—indeed, on a completely static view, virtually impossible. This is why there is an equilibrium at A—as Mr. Harrod, for different (or for no!) reasons, correctly affirms.

On the other hand, the equilibrium of 'accepted doctrine' at B is necessarily unstable. For while profits are 'normal', the price is in excess of the cost of efficient production, i.e. production at some scale beyond OM in Fig. 1, and the prospects of forcing an entry are thereby enhanced. New entrants will thus be attracted and a struggle for the custom of the market will necessarily ensue. We return to this point in § 5.0 below.

4.2. Next, is it at all reasonable to assume that while the entrepreneur acts oligopolistically with respect to new competition, he yet fails to have

regard to his established competitors? Certainly it would be difficult to locate such an entrepreneur in practice. On the contrary, firms resort constantly to a variety of measures to gain information about what their established rivals are doing, and they act on this knowledge.[1]

Now oligopolistic behaviour in this sense—as well as in the sense of having regard to potential competitors—is integral to Mr. Andrews's system. In excluding this aspect of entrepreneurial behaviour from view, Mr. Harrod's analysis is altogether too 'full-costy'. He envisages as a *determinant* of the firm's equilibrium a curve of long-period average (and marginal) cost which in fact is fully determinate only *in* the equilibrium. This procedure may doubtless be defended on the grounds of simplicity. But the simplification is rather dearly bought, for it serves also to obscure another important tendency in the pricing of manufactured commodities, namely, that the price tends to be equal to the average inclusive cost of the most efficient of the extant firms in equilibrium.

These points are important, and a fuller discussion of them will serve to clarify Mr. Andrews's approach to the theory of price determination in manufacturing industry.

4.2.1. For this purpose we may argue in terms of the simplified model of Mr. Andrews's system, which I have analysed more fully elsewhere.[2] The elements of this model are set out in the first section of the following.

(1) The model envisages a 'manufactured' intermediate product. In this sphere of industry (as indicated earlier) a relatively stable channelling of the aggregate demand as between the particular products of the several manufacturers tends to emerge in terms of 'business goodwill'. The buyers become attached to particular sellers in terms of the confidence given by previous custom; and the sellers seek to cement these ties, because in the absence of an organized market the necessary basis of efficient manufacture is a stable and continuing demand for the output of the firm just such as the regular custom of a definite group of buyers provides.

Now the existence of this 'business goodwill' is not incompatible with a situation in which the products of the various manufacturers are in fact (i.e. objectively) identical in all respects, save for the particular identification of each.[3] For while in these circumstances the initial choice of a buyer may have no particular rational basis, subsequently from his point of view the product and associated services—including the regularity of the supply—of the manufacturer from whom he buys has been, and is

[1] See, for example, J. K. Galbraith, *American Capitalism* (London, 1952), p. 15 *et passim*.

[2] See my article in *Economic Record*, referred to above.

[3] i.e. the products are physically standardized; similar facilities as to credit are offered, &c.

being, proven; but not so the products and associated services of the other manufacturers.

The working assumption of this simplified model will be, therefore, that the products of the several manufacturers are in fact identical in all respects. In practice there is a definite tendency in this direction within any sub-class of a broad commodity group; because any variation in product design, sales service, and so on, which one manufacturer may introduce and which *does* prove successful (i.e. wins over demand at the maintained price), must needs be imitated by his rivals. Our procedure thus amounts to this: that we abstract, temporarily, from the constant thrust and parry in terms of (minor) innovations in the design and quality of the product, which is the dynamic essence of competition in practice, in order to concentrate attention on the (static) long-run equilibrium tendency.

Now in the situation envisaged, *the goodwill of all necessarily limits the sales of each*; and this may indeed be represented by a downsloping 'demand curve'. (This will intersect a horizontal line, drawn through the prevailing price, at the firm's current volume of sales—see the *d* curves in Fig. 2 below.) But this is not an operative sales curve for the firm, in the sense of Professor Chamberlin's 'particular market' (*dd'*) curve. That is a stable function, showing (for a higher price) the volume of demand the seller would *continue* to command if his rivals maintained their prices unchanged and no new competitors entered the industry. In the present model, no such curve exists, because it is a basic assumption of the abstract Normal Cost theory that 'the buyers are rational in the sense that they will not continue to pay more for the same quality of product than they need pay elsewhere'. It follows from this assumption that, after some interval, the demand of the firm at the higher price would recede, and in the limit to zero.

As applied to the case before us, this abstract assumption is amply justified. The buyer is himself an entrepreneur. He may pay a higher price for a time—because of the risks and inconvenience of a change in his source of supply (in face of price rise which may soon become general), and because of the personal element in the goodwill relation. But these factors would not long outweigh the pecuniary advantages of buying more cheaply elsewhere; the buyer has his own competitive position to consider. After some interval (and then with increasing momentum), therefore, the customers of the firm would change over to alternative suppliers.

This would amount to a shift in the initial 'demand curve' to the left, *not* as the result of the entry of new competitors. Moreover the original position could not readily be restored; because the loss of demand would be the loss of 'goodwill' (or the market connexion of the firm), which has

been—and could again be—acquired only with difficulty and over a long period of time.

The crucial question, therefore, is the length of this 'demand-transfer period', i.e. the average period during which the buyers will change over to alternative suppliers, relative to the period during which the firm expects to recoup the capital outlay on the plant. The evidence would suggest that this is relatively short. Normally, therefore, the individual producer will not prejudice the security of his investment by charging a price higher than other producers of the same quality of product.

The situation differs structurally from the case of Perfect Competition, but is no less 'competitive'. Perfect Competition envisages an *instantaneous* transfer of demand from the higher-priced to the lower-priced sources of supply: in point of fact, it is scarcely within the purview of that theory to consider a price higher than the going market price at all. The element of 'goodwill', on the other hand, *does* enable a temporary divergence of price; but unlike the case of Perfect Competition, the subsequent loss of demand is not readily regained. It is most important that this latter consideration should be set against the instantaneous substitution of the Perfect Competition case, in passing judgement on the relative 'competitiveness' of the two situations.

The individual producer, then, will not maintain a price higher than other producers of the same quality of product. Similarly a price cut by one producer will be matched after some interval by all.

It follows that in the long run and notwithstanding the element of 'goodwill', *one price will prevail for the products of the several manufacturers.* At this price, the aggregate custom of the market will be divided between the manufacturers in terms of the 'goodwill' each commands.

(2) The question now is: How is the 'normal' price determined? Referring to Fig. 2, let us suppose for simplicity that there are but three firms in the industry whose long-run average cost curves are as shown in the left-hand figure. These are of the same general shape as envisaged also by Mr. Harrod; but the curve of each firm varies at any given level of output in absolute distance from the x axis according to the firm's 'efficiency'.

These curves may be portrayed as fully determinate *ex ante* if we assume that net profit is included in each (and for every scale of each) at the same rate per cent., say 10 per cent., on the capital invested. This is not to say that each, or even any, firm will *realize* this rate in equilibrium. It is assumed, however, (without subtracting from the generality of the argument), that the rate in question is *of the order* that each entrepreneur would consider an adequate return on his capital.

We may suppose that in the early stages of the development of the industry, with each of many small firms producing at a scale less than Ox

in the left-hand figure, the price was high; but subsequently, with the growth in the aggregate demand and in the size of the typical firm, it has fallen—as indicated by the L curve to the left of D_t in the right-hand figure. At time t (our initial position), the aggregate demand curve has stabilized at D_t; the prevailing price is OP_t; each firm's share of the market is assumed to be the same (Ox units), and is indicated by the d_t curve, common to all; and the plant curves of the three firms, at the assumed rate of net profit, are shown (ac_t).[1]

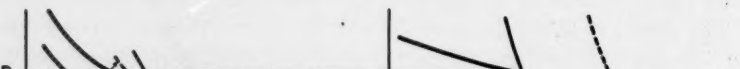


FIG. 2.

Consider now the position of the most efficient firm. In the first place, the *expectation* of the entrepreneur is that, if he could but expand sales, then with the appropriate reorganization of the plant the firm could produce at a lower cost per unit. For his expectation is guided only by the continuously falling technical costs of production,[2] since he will not expect himself (or the management generally) to become less efficient just because the firm grows—albeit in the event, this may turn out to be so. To increase sales is thus of paramount concern to the entrepreneur; to this end, the ample margin of the prevailing price above the calculated costs of the firm provides plenty of room for manœuvre in an attempt to eliminate the high-cost firm 3 (representative of the inefficient generally).

Second, the entrepreneur cannot but regard the high profits of the firm at the price OP_t—which are considerably in excess of the level envisaged in his initial calculation—as precarious. This high profit is due to the firm's superior efficiency, and this—to use a constantly reiterated phrase of Mr. Harrod's—is the result of its 'possessing an advantage peculiar to itself' (op. cit., pp. 143, 151, 161, &c.). But this 'advantage' is peculiar to the firm *only as amongst the existing firms*; the possibility remains that some firm among its *potential* rivals may be able to appropriate a similar 'advantage'—to emulate the firm's superior efficiency.

[1] These curves intersect the *LRAC* curve because of the *planned* reserve of capacity (for breakdowns; to meet peak demand; &c.). See Andrews, op. cit., pp. 87–95.

[2] See Andrews, op. cit., pp. 113–23, 136–7.

The essential point is that there would seem to be no factor of production in manufacturing industry analogous to land in agriculture. Fertile land, in close proximity to the marketing centre, is limited in supply, and confers an advantage in production which no new entrant can emulate (thus the intra-marginal firm of the Perfect Competition model). In manufacture, on the other hand, the basic factor machines (and machines of whatever efficiency) can be duplicated without limit. The factor of ultimate business capacity is somewhat different in that it would be incorrect to assume that all prospective new entrants should possess—or be able to command—entrepreneurial skill of the highest order. Yet it is surely correct to assume that *some* at least of potential new entrants—particularly the more efficient of established firms in adjacent lines of production—should possess this skill; i.e. that the range and variation of abilities among potential rivals is roughly similar to that of the existing firms. In a word, the skilled entrepreneur among potential rivals may challenge even the ablest of established producers—if the prospects are sufficiently attractive.

It follows that the price OP_t is not an equilibrium one. For either the most efficient (at least) of the established firms, aware of the possibilities of new and efficient competition, will cut the price; or, if it should not do so, this new competition will in fact materialize—with the same result, as far as the price is concerned. In Fig. 2 equilibrium is pictured as obtaining at the price OP. In general terms, the conditions of the long-period equilibrium are (1) that the aggregate demand at the price set by the most efficient is within the capacity of the extant firms, and (2) that there is no tendency for firms to enter or leave the industry, i.e. each firm now remaining in the industry considers the profits it is earning a sufficient return on the capital invested, and no potential new entrant deems the price (and prevailing profits) such as to justify entering the industry.[1]

[1] In Fig. 2 it is assumed that firm 1 did cut the price; that at this price the profits (or rather, losses) of firm 3 were such that it has been liquidated; and that firms 1 and 2 have appropriated the additional custom xx_1 and xx_2—market reaction 'demand' curves d_1 and d_2—respectively.

If firm 1 had not initiated the price reduction, then the market shares and the number and identity of the firms in equilibrium may well have been different; these thus depend on the path to equilibrium. This is also true of the Perfect and Imperfect Competition models; one might say that the number of firms is determinate in the *neighbourhood* of equilibrium. It will be argued in § 5.1 below that the tendency will be for each firm that remains in equilibrium to be producing at, or beyond, the scale Ox; this scale of production relative to the size of the aggregate demand, will thus set an upper limit to the number of firms in equilibrium. Thus, too, the *price* itself is not dependent on the path to equilibrium.

Finally, the equilibrium is static in that it is assumed (*a*) that the state of technical knowledge is given, and (*b*) that the structure of relative efficiencies of the firms in equilibrium remains constant. Variations in the latter in practice give rise to that 'constant passing to and fro of industrial leadership' which is the dynamic essence of competition in practice. (See Andrews, 'Industrial Analysis in Economics', *Oxford Studies in the Price Mechanism*, pp. 145, 171, *et passim*.)

4.2.2. Three points should now be noticed.

(A) The foregoing conditions *relating to the market as a whole*, define the equilibrium price. The pricing problem of the *individual firm* is then seen for what it is, namely, to 'find' this price—the 'right' price for its product. It will do this in the only manner possible: first, on the basis of its own costing (this will often take the form of applying some costing 'rule of thumb'); and overriding this, second, by looking to what its rivals are doing in both its own and adjacent lines of production (potential competitors), and adjusting its policy accordingly.[1]

The firm must necessarily proceed in both these ways, because its own projected 'cost' of production can be, in the nature of the case and however elaborate its cost-accounting, an estimate only. This is so because the appropriate (equilibrium) rate of net profit of the firm is finally determinate only *in* the overall equilibrium. (The equilibrium profits, in general different for each firm, may be defined as 'normal', analogously to the textbook model of Perfect Competition.)

Thus oligopolistic behaviour is inherent in Mr. Andrews's approach, and this irrespective of whether the number of competitors is large or small; because in either case, the individual entrepreneur's own costing estimate is equally fallible. This is of the essence of Mr. Andrews's defection from the 'full cost principle'.

(B) So much for the pricing policy of the firm. As to output, the best that the firm can do is to supply the custom it commands at the equilibrium price to the full; striving always to increase its share of the aggregate demand, because of the expected fall of long-run cost with increased scale.

The L curve beyond E in Fig. 2 is thus not a (long-period) supply curve in the traditional sense. In the equilibrium at E, each firm *could* supply more, and *would* do so, and *at an unchanged price* in both the short and the long periods, if the additional demand were there; so that an increase in the aggregate demand, to D' in the figure, will in general be wholly supplied by the *existing* firms, and the new long-period equilibrium at E' will be attained without any rise in the price.

This result follows, because in the short (or transition) period the firm will not risk losing prospective additional customers, or endangering its present goodwill, by charging a higher price now in accordance with any rise in short-period marginal cost. It will, if necessary, resort to shift and overtime working.[2] In the long period it will extend its plant, the *LRAC*

[1] A simple illustration from the model: if initially firm 3 in the light of $LRAC_3$ had judged OP_4 as about 'right', this judgement (and the costing method) would necessarily have been revised as firm 1 cut the price (or, alternatively, as the threat of new competition became actual).

[2] In Mr. Andrews's view, of course, the rise in short-period marginal cost immediately

curve (it is now generally agreed) being approximately horizontal over a very large range of output.[1]

The result that the rise of short-period marginal cost for the given plant is in the main irrelevant, may be expressed by saying that for the individual firm there is an equilibrium of *price*, but not a determinate *price-output* equilibrium in the usual sense. In a word: the 'right' (equilibrium) price is independent of the planned, or achieved, output of the firm. Hence the invariance of the 'oncost percentage' for moderate fluctuations in output about the planned throughput of the firm; no extension of the so-called 'ritual' (i.e. the firm's costing rule), such as Mr. Harrod has suggested (p. 168), is needed to explain this.

In Mr. Harrod's analysis, on the other hand, short-period marginal cost retains its traditional role; its point of intersection with the virtually horizontal 'long-period marginal revenue curve' determines short-period output (p. 150). But in the event of an increase in demand beyond this point, what of the unsatisfied demand at the maintained price?

At a later stage, Mr. Harrod refers to the 'waiting list' (pp. 167–8), characteristic of the post-war period of full employment. As he himself points out, however, 'fear of competition is a two-edged argument in this connexion. High prices might stimulate competition; but so also might the waiting lists, if rivals could acquire the necessary materials and machines' (p. 168). The conditional clause is significant. The post-war period was one of universal shortage of resources. The established firm, therefore, could introduce a waiting list with some confidence, because if such a firm were unable to obtain machines and men to expand its output, the would-be new entrant would find himself in a similar pass.[2] In normal conditions, however, the case is different; so that in the event of a structural (i.e. non-cyclical or non-inflationary) increase in demand, the typical manufacturing firm with a regard for potential competition is not likely to resort to the waiting list. We come back to Mr. Andrews's position: that in view of the possibilities of overtime working, sub-contracting, &c., in the short run, and the overriding importance of maintaining and extending the goodwill of the firm, rising short-period marginal cost is in the main irrelevant.

beyond the normal throughput of the existing plant is not so steep as usually envisaged, because of the planned reserve of capacity previously referred to.

[1] The complete contrast is with the Perfect Competition model of constant returns (*LRAC* curve of the firm U-shaped), where in the event of an increase in demand the price will rise initially, yielding excess profits, while finally the whole of the increased demand will be supplied by *new* firms at the previous price.

[2] If this view is correct, it becomes necessary to explain the price restraint (and waiting list). This was mainly a matter of complying with government appeals in the name of internal stability and external solvency, at a time of national crisis. (It should be added that the possibilities of overtime working, &c., were limited, because the period of shortage was prolonged and plants could not be overworked indefinitely.)

(c) Finally, Mr. Harrod is, of course, correct in representing the price as equal to the firm's average inclusive cost of production. But this is no more than a definitional identity; because the average cost is defined to include normal profit, which (per unit) is equal to the difference between the (equilibrium) price and the average of all other costs.

What is important is that the price tends to equal the average inclusive cost of the *most efficient* of the extant firms in equilibrium. Mr. Harrod's procedure obscures this result. He affirms, repeatedly, that '. . . [the entrepreneur] will be able to quote a price that covers his overhead costs and gives a normal profit—[*and*] *more if he has superior efficiency* . . .'.[1] Now if this means merely that the more efficient firms will be able to earn a higher margin of net profit over and above the average of other costs, as compared with the less efficient, then this is undoubtedly correct. But his oft-repeated turn of phrase—that this higher unit net profit is appropriated because such a firm 'possesses an advantage peculiar to itself'—is suggestive of the traditional intra-marginal firm and, by association, of the result that the price is determined by the cost of production of the marginal (or *least efficient*) firm. That would indeed be unfortunate.

We have argued that in manufacturing industry, first, there is normally no advantage in production which an efficient new entrant might not emulate, and, second, there is virtually no limit from the production (cost) side to the continued expansion of an established firm (horizontal scale curve). This last means that the more efficient of the firms (in particular) could, and would, supply the whole market—were they but able to wrest from their less efficient competitors the custom these latter command. The production of the latter, therefore, is in no way 'necessary if the aggregate demand is to be met at the price', as with the marginal firm(s) in Perfect Competition so; that the price bears no necessary relation to the supply price (cost) of this output. On the contrary, the most efficient firm will set the price (and the inefficient may hang on if they can); because no firm will tend (or continue) to quote lower than this firm, while it will itself remain subject to the threat of similarly efficient competition from without.

The Doctrine of Excess Capacity

5.0. The ultimate purpose of Mr. Harrod's essay is to refute the doctrine that there is necessarily a tendency to 'excess capacity' in conditions of 'Imperfect Competition with free entry'.

Thus in the equilibrium of 'accepted doctrine' at B in Fig. 1, not only is the installed plant (ac curve) being operated at less than the conventional optimum (point of minimum short-period average cost), but also the plant

[1] *Essays*, p. 161 (my italics). See also pp. 143 (second par.), 151 (second par.), 159, &c.

itself is of less than optimum capacity (scale). These are the short-run and long-run aspects of the excess capacity doctrine, respectively. The former represents the condition where each firm, because its throughput is too small, is loading too much overhead into the unit price; this is the excess capacity problem proper, which Mr. Harrod has presumably in mind (p. 146). The long-run aspect—which is more appropriately termed the 'problem of unexhausted economies of scale'[1]—is that with which Professor Chamberlin is concerned in *The Theory of Monopolistic Competition*.[2]

Now Mr. Harrod denies the validity of the equilibrium at *B*, and argues to the 'equilibrium' at *A* in Fig. 1. In this there is no excess capacity in either of the two senses just indicated; accordingly (he contends) the doctrine of excess capacity must be abandoned.

5.1. Thus far, however, the argument against excess capacity is not conclusive. In the absence of reasoning to the contrary (and Mr. Harrod does not supply this), the *d* curve might equally well have intersected the downward-sloping section of the *LRAC* curve in Fig. 1, in which event—and even though, in Mr. Harrod's view, the entrepreneur will charge a price equal to the average cost at this point—there will be excess capacity in both senses. Mr. Harrod recognizes this possibility (p. 181).

It is necessary, therefore, to ask, why should the market of the individual firm be so restricted as in this latter case, in the long-period equilibrium position? The same question arises when we attempt to pass beyond the mere geometry of the 'tangency' equilibrium of 'accepted doctrine', and consider its realistic content.

The answer in the case of 'accepted doctrine' is to be sought, first, in the assumption that entry is not merely free, but also easy. This implies that custom is easily attracted; so that, as the existing firms equate short-period cost and revenue and new entrants are attracted, each firm—both the old and the new—will be very much an 'average' firm in the sense that its demand will necessarily be limited to an *n*th part of the total demand at the prevailing price (where *n* is the number of firms in the neighbourhood of equilibrium). This surely is what is meant when excess capacity is attributed by Mrs. Robinson to 'the imperfection of markets combined with ease of entry'.

But this is not enough. With each individual market so restricted as in the equilibrium at *B* in Fig. 1, it would require a battery of unrealistic assumptions if the position were to remain so in the long run; that is to say,

[1] See T. Wilson, 'Inadequacy of the Theory of the Firm as a Branch of Welfare Economics', *O.E.P.* Feb. 1952, *passim*.

[2] See ch. v, § 5, pp. 104–9 (4th ed.). The *PP'* curve of Professor Chamberlin's diagrams is clearly a long-run cost curve (p. 75).

if—in view of the substantial unexploited economies of scale beyond OM_4 in the figure, which would enable a lower price (or greater profits) if the necessary throughput were achieved—the more efficient, the more affluent and/or the more enterprising among either existing or potential competitors were not to initiate a price war to drive out the less-favoured firms, and so expand their own scales of production. That this should not occur would require that 'the firms must be equally matched in financial strength if the outcome of a price war [were] to be in doubt'; that 'they must be more or less equally efficient for, if this were not the case, some would earn larger profits at the stable price and gain financial dominance'; that the entrepreneurs are all equally lacking in vision, and evenly matched in business acumen and daring; and so on.[1] These assumptions are inappropriate to a general theory; the 'equilibrium' at B, therefore, cannot represent a *stable* long-period situation.

It is for these same reasons that each firm in Fig. 2 above is depicted as producing at, or beyond, the 'practicable optimum' scale of production Ox, i.e. the minimum scale of operations necessary to achieve the major economies of production available at the existing stage of technical development. The extent of the 'goodwill' the firm commands in the market effectively limits its sales in the short run (i.e. at any given point of time). But so long as there are significant economies to be had by larger-scale production, this barrier can, and will, be surmounted—by the voluntary merger or by aggressive price (and non-price) competition to eliminate less able competitors and appropriate the custom these command. This sort of consolidation process, as we have argued previously, typically follows after the initial 'mushroom' of firms in the course of the development of a new industry.

Thus excess capacity in the sense of substantial unexploited economies of scale will not normally obtain. And to return then to Mr. Harrod's situation where the market of the firm is 'more restricted' than at A: this could arise in the long run only in the event that the 'practicable optimum' were so large in relation to the *aggregate* demand for the product that not even one producer, when he—the most efficient and boldest of all—has appropriated to himself the great bulk of the demand, can attain the necessary scale of production. It is a question of fact whether this is, or is not, a frequent situation in practice; the writer—and I think Mr. Harrod—would judge that it is not.

5.2. It will be noted, finally, that the 'planned reserve capacity' in Mr. Andrews's system amounts to short-period excess capacity, in established

[1] T. Wilson, op. cit., pp. 30–31.

terminology. But first, inasmuch as it is *planned*, the case is different from the (short-period) excess capacity of 'accepted doctrine', which is *imposed* on the entrepreneur by the overall equilibrium. Second, it often consists in practice (as Mr. Andrews has himself argued) of equipment which has been replaced in the normal course of events by new machinery, and has been withheld from the scrap-heap for just these purposes (i.e. as a reserve in the event of breakdowns, unexpected increases in demand, &c.).

When this has been said, however, it remains true that this sort of excess capacity is a necessary cost of flexibility in the quasi-free enterprise system of our times.

'Free Competition'

6.0. One further point. Mr. Harrod suggests that economists might do well to appropriate the popular term 'free competition' to describe the case of Imperfect Competition under discussion together with that of Perfect Competition.

Some such modification of the traditional nomenclature is, I think, overdue. The traditional adjectives 'perfect' and 'imperfect', as applied to competition are, when juxtaposed, highly emotive and persuasive. To the extent that it was contended that Imperfect Competition meant prices higher and outputs more restricted, as compared with Perfect Competition, there was some justification for the suggestive content of this nomenclature. But these theses are now in abeyance. (Capitals are used throughout this paper to emphasize that 'Perfect' and 'Imperfect Competition' are technical terms, not necessarily literally descriptive.)

Mr. Harrod's suggestion, however, is itself open to objection on the grounds of emotive content. 'Free competition' is suggestive of all the real (and imagined) virtues of 'free enterprise' and nineteenth-century *laissez-faire*. This sort of reference is also likely only to cloud discussion.

It would seem that the safest course would be to seek some entirely neutral terminology. Thus the two broad cases envisaged in this discussion —'Perfect Competition', and 'Imperfect Competition with free entry'— might be designated, barely, as 'competition of the first kind' and 'competition of the second kind'. The first of these abstract analyses would refer to that 'sizeable part of total production—most food and raw materials in the world' (Harrod, p. 175); the second, to the ordinary run of manufacturing and commercial business. (Of course there would be many variants within each of these two categories, since in practice every industry is in some respects a special case. And with the future development of the science, an extension of the basic categories would doubtless emerge.) The present suggestion would at least have the merit of putting

an end to that misunderstanding between economists and businessmen which arises from a concept of '*perfect* competition' in which all that the latter would understand by competition is virtually absent.

THE UNIVERSITY,
SYDNEY,
AUSTRALIA.

OXFORD ECONOMIC PAPERS

Volume 6. September 1954. Number 3

CONTENTS

Period Analysis and Inventory Cycles. By R. NURKSE . . . 203

The Price of Fuel. By D. L. MUNBY 226

The Dollar Crisis Revisited. By T. BALOGH 243

Investment Decisions, Trade Cycle, and Trend. By A. J. YOUNGSON 285

Some Elementary Theory About Accumulation. By D. B. BUTT . 306

ECONOMICA

Published Quarterly

ECONOMICA is published quarterly by the London School of Economics and Political Science, in February, May, August, and November, and is devoted to research in the fields of Economics, Economic History and Statistics. The Journal is under the direction of an editorial board composed of Professor R. G. D. Allen, Professor T. S. Ashton (Acting Editor), Professor W. T. Baxter, Professor E. H. Phelps Brown, Sir A. M. Carr-Saunders, Professor R. S. Edwards, Professor J. E. Meade, Professor F. W. Paish, Professor Sir Arnold Plant, Professor Lionel Robbins, Professor R. S. Sayers (Acting Editor), with Mr. A. T. Peacock as Assistant Editor.

34th Year New Series Volume XXI No. 84 November 1954

CONTENTS

Economics of Dependence (1952–82) *F. W. Paish* and *A. T. Peacock*
Consistency and Consolidation in the Theory of Interest *Ralph Turvey*
Industrial Productivity Growth in Europe and in the U.S. *A. Maddison*
The Origin of Ricardo's Theory of Profits *G. S. L. Tucker*
Non-Traded Goods and International Factor Price Equalization *E. Victor Morgan* and *G. L. Rees*

BOOK REVIEWS BOOKS RECEIVED

The price of ECONOMICA (New Series) is 8*s.* 6*d.* a number, or £1. 5*s.* per annum, post free. Copies may be obtained on order from the Publications Department of the London School of Economics, Houghton Street, Aldwych, London, W.C. 2, to which address all subscriptions and business communications should be sent.

THE ECONOMIC JOURNAL

The Quarterly Journal of the Royal Economic Society

DECEMBER 1954

UTILITY AND ALL WHAT? — Sir Dennis Robertson

SOCIAL ACCOUNTING IN RELATION TO ECONOMIC THEORY — D. K. Burdett

THE REDUCTION OF FLUCTUATIONS IN THE INCOMES OF PRIMARY PRODUCERS: A CRITICAL COMMENT — M. Friedman

THE REDUCTION OF FLUCTUATIONS IN THE INCOMES OF PRIMARY PRODUCERS FURTHER CONSIDERED — P. T. Bauer and F. W. Paish

FLUCTUATIONS IN INCOMES OF PRIMARY PRODUCERS: FURTHER COMMENT — B. M. Niculescu

THE PROBLEM OF INTERNATIONAL INVESTMENT TODAY IN THE LIGHT OF NINETEENTH-CENTURY EXPERIENCE — R. Nurkse

THE POSITION OF ECONOMICS AND ECONOMISTS IN THE GOVERNMENT MACHINE — R. L. Marris

STATISTICAL INFERENCE IN DYNAMIC ECONOMIC MODELS — M. R. Fisher

MACMILLAN & CO., LTD., LONDON, W.C.2

Application for Fellowship to
Royal Economic Society, 21 Bentinck Street, London, W.C. 2
Annual Subscription £1 10 0 **Life Composition £22 10 0**

SCOTTISH JOURNAL OF POLITICAL ECONOMY

The Journal of the Scottish Economic Society

Volume II *February 1955* *Number 1*

1. THE PRESENT COMPLEXITY OF WAGE PAYMENTS — D. J. ROBERTSON
2. THE AMERICAN ORIGINS OF A SCOTTISH INDUSTRY . . . . . — W. WOODRUFF
3. BUSINESS MEN ON BUSINESS IDEAS . . — G. L. S. SHACKLE
4. ECONOMIC WELFARE AND MONEY INCOMES IN THE HIGHLANDS, 1750–1850 . . . — M. GRAY
5. RECENT ECONOMIC TRENDS

Scottish Journal of Political Economy is edited by Professor A. K. Cairncross, is published in February, June, and October, and is free to members of the Scottish Economic Society. Applications for membership should be made to the Treasurer, Scottish Economic Society, 142 St. Vincent Street, Glasgow, C. 2, enclosing the annual subscription of one guinea. The Journal may also be obtained without membership of the Society by ordering through any bookseller or direct from the Publisher, Oliver and Boyd, Tweeddale Court, Edinburgh. For Volume I the price of single copies is 7*s.* 6*d.* each, and from Volume II onwards 10*s.* each.

PRINTED IN GREAT BRITAIN, AT THE UNIVERSITY PRESS, OXFORD
BY CHARLES BATEY, PRINTER TO THE UNIVERSITY